

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS NATURAIS E EXATAS (CCNE)
PROGRAMA DE PÓS-GRADUAÇÃO EM BIODIVERSIDADE ANIMAL

Marcos Trindade da Rosa

**DESCRIÇÃO DO GENOMA, TRANSCRIPTOMA E MICROBIOMA DE
STENOSTOMUM LEUCOPS (PLATYHELMINTHES, CATENULIDA), COM
ENFOQUE NA FILOGENIA, REGENERAÇÃO, PLASTICIDADE ECOLÓGICA E
ELEMENTOS TRANSPONÍVEIS.**

Santa Maria, RS

2022

Marcos Trindade da Rosa

**Descrição do genoma, transcriptoma e microbioma de *Stenostomum leucops*
(Platyhelminthes, Catenulida), com enfoque na filogenia, regeneração, plasticidade
ecológica e elementos transponíveis.**

Orientador: Elgion Lucio da Silva Loreto

Tese apresentada ao Programa de Pós-Graduação em
**Biodiversidade Animal, da Universidade Federal de
Santa Maria (UFSM, RS), como requisito parcial
para obtenção do título de **Doutor em
Biodiversidade Animal.****

Santa Maria, RS, Brasil

2022

Trindade da Rosa, Marcos

DESCRIÇÃO DO GENOMA, TRANSCRIPTOMA E MICROBIOMA DE STENOSTOMUM LEUCOPS (PLATYHELMINTHES, CATENULIDA), COM ENFOQUE NA FILOGENIA, REGENERAÇÃO, PLASTICIDADE ECOLÓGICA E ELEMENTOS TRANSPONÍVEIS. / Marcos Trindade da Rosa.- 2022.

107 p.; 30 cm

Orientador: Élgion Lucio Loreto da Silva
Tese (doutorado) - Universidade Federal de Santa Maria, Centro de Ciências Naturais e Exatas, Programa de Pós-Graduação em Biodiversidade Animal, RS, 2022

1. Stenostomum leucops 2. Filogenia 3. Microbioma 4. Genoma 5. Transcriptoma I. Lucio Loreto da Silva, Élgion II. Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB 10/1728.

Declaro, MARCOS TRINDADE DA ROSA, para os devidos fins e sob as penas da lei, que a pesquisa constante neste trabalho de conclusão de curso (Tese) foi por mim elaborada e que as informações necessárias objeto de consulta em literatura e outras fontes estão devidamente referenciadas. Declaro, ainda, que este trabalho ou parte dele não foi apresentado anteriormente para obtenção de qualquer outro grau acadêmico, estando ciente de que a inveracidade da presente declaração poderá resultar na anulação da titulação pela Universidade, entre outras consequências legais.

Marcos Trindade da Rosa

Descrição do genoma, transcriptoma e microbioma de *Stenostomum leucops* (Platyhelminthes, Catenulida), com enfoque na filogenia, regeneração, plasticidade ecológica e elementos transponíveis.

Tese apresentada ao Programa de Pós-Graduação em **Biodiversidade Animal**, da **Universidade Federal de Santa Maria (UFSM, RS)**, como requisito parcial para obtenção do título de **Doutor em Biodiversidade Animal**.

Aprovado em 31 de janeiro de 2022:

**Élgion Lúcio da Silva Loreto Dr. (UFSM) - Videoconferência
(Presidente/Orientador)**

Daniel Ângelo Sganzerla Graichen Dr. (UFSM) - Videoconferência

Lizandra Jaqueline Robe Dra. (UFSM) - Videoconferência

Gabriel da Luz Wallau Dr. (FIOCRUZ – IAM) - Videoconferência

Karen Luisa Haag Dra. (UFRGS) - Videoconferência

Santa Maria, RS, Brasil

2022

NUP: 23081.009506/2022-41

Prioridade: Normal

Homologação de ata de banca de defesa de pós-graduação

134.332 - Bancas examinadoras: indicação e atuação

COMPONENTE

Ordem	Descrição	Nome do arquivo
2	Parecer de defesa de dissertação/tese (134.332)	folha de aprovação (1).pdf

Assinaturas

07/02/2022 17:00:43

LIZANDRA JAQUELINE ROBE (PROFESSOR DO MAGISTÉRIO SUPERIOR)

02.75.00.00.0.0 - DEPARTAMENTO DE ECOLOGIA E EVOLUÇÃO - DEE

07/02/2022 17:32:58

ELGION LUCIO DA SILVA LORETO (PROFESSOR DO MAGISTÉRIO SUPERIOR)

02.71.00.00.0.0 - DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR - BBM

08/02/2022 16:50:11

DANIEL ANGELO SGANZERLA GRAICHEN (PROFESSOR DO MAGISTÉRIO SUPERIOR)

32.03.00.00.0.0 - VICE-DIREÇÃO DO CAMPUS DE PALMEIRA DAS MISSÕES - VD-PM

08/02/2022 18:14:18

Karen Luisa Haag (Pessoa Física)

Usuário Externo (456.***.***.**))

09/02/2022 05:47:01

GABRIEL DA LUZ WALLAU (Pessoa Física)

Usuário Externo (014.***.***.**))



Código Verificador: 1141444

Código CRC: b34c8b68

Consulte em: <https://portal.ufsm.br/documentos/publico/autenticacao/assinaturas.html>



“É claro que os biólogos moleculares e celulares estão experimentando anseios semelhantes, querendo provar as delícias científicas de organismos não-modelo.”

William Sullivan, 2015.

Agradecimentos

Agradeço primeiramente ao meu orientador, professor Elgion Lucio da Silva Loreto, por ter me proporcionado esta oportunidade. Por ter aceito um verme entre moscas, com suas peculiaridades e desafios. Disponibilizado tempo e recursos com minha formação, e me aceitado como parte da família Labdros. Alguém que tenho um respeito e carinho gigantesco, por ter sido meu pai científico, além de amigo nestes quase 12 anos.

Agradecer aos colegas, que compartilharam conhecimento, desejos, alegrias, medos e tristezas. Amizades que são além das quatro paredes de um laboratório, e até mesmo a instituição UFSM. Muitos nomes e personalidades passaram nestes anos, pessoas que sempre estarão em minhas memórias. Aquele momento de aprendizado: Gabinho ensinado a fazer meio de *Drosophila*, Ronaldinho a diluir *primers*, Mauro me apresentando um alinhamento, Paloma fazer um PCR, Junges apresentando a diversidade de *Drosophila*, as dores de cabeça nas discussões de bioinformática com o Zé, Daniel e Tiago.... Momentos de descontração em nossos churrascos, e sim dá para temperar carne com tempero verde “ervas finas”. Há muitos nomes não citados, que são responsáveis por todo este e demais trabalhos desenvolvidos.

Agradecer a pessoas que foram muito importantes na minha graduação, que não são simples colegas, mas amigos que levo até hoje nos laços afetivos. A Cami (Camila de Moura), por ter me aguentado seis anos nas aulas e laboratório. A Li (Livia Bataioli), que além de momentos da biologia, me arrastava para os campos da herpetologia. A Lu “canhinha” (Luiza Flores), por ter insistido em me tornar mais acessível as pessoas e amizades, além de me arrastar nos campos de ornitologia “sobe morro, desce morro, as cinco horas da manhã”. Pessoas que compartilhamos sonhos, medos, tristezas, felicidade e ainda espero continuar compartilhando por muitos anos.

Agradecer ao financiamento da pesquisa pela CAPES, ter acreditado no projeto e minha capacidade em executá-lo.

Agradecer meus pais, por insistirem que eu deveria voltar aos estudos. Assim me permitindo ter tido todas estas experiências e conhecido pessoas incríveis. Estes meus pilares, que junto ao professor Elgion sempre me guiam. Permitindo-me conhecer a biologia de uma maneira inimaginável, e quanto este micro-platyhelminthe me cativou.

Agradecer também a nova geração Labdros, que se permitiram conhecer o *S. leucops*, e estão dispostos a continuar a mergulhar no mundo dos micro-platyhelminthes.

Resumo

Descrição do genoma, transcriptoma e microbioma de *Stenostomum leucops* (Platyhelminthes, Catenulida), com enfoque na filogenia, regeneração, plasticidade ecológica e elementos transponíveis.

AUTOR: Marcos Trindade da Rosa

ORIENTADOR: Elgion Lucio da Silva Loreto

Platyhelminthes é o filo com maior radiação adaptativa dos invertebrados, e está dividido atualmente em cinco classes: Catenulida, Rhabditophora, Monogenea, Cestoda e Trematoda. As análises filogenéticas, empregando 12 genes mitocondriais de 165 espécies de Platyhelminthes, posiciona Catenulida como grupo irmão de Rhabditophora. Sequenciamento *long reads* revelaram, em algumas espécies de Platyhelminthes, que os mitogenomas podem ser maiores do que os apresentados por sequenciamentos do “*short reads*”. *S. leucops* apresenta poucos tecidos diferenciados e é rico em células-tronco (*stem cells*), responsáveis por sua regeneração e reprodução. A literatura apresenta descrições contraditórias sobre aspectos regenerativos de *S. leucops* na formação de zooide, quando se remove a parte cefálica anterior e o corpo posterior. Os experimentos refeitos por nós, constataram que todas as maneiras de regeneração descritas na literatura, podem ser obtidas. O microbioma de *S. leucops* parece ir além de simples alimento para o verme, pois este têm capacidade de expressar DNA contido em bactérias (plasmídeo com marcador *gfp*). *S. leucops* também tem capacidade de utilizar “DNA livre” presente em seu meio, mesmo em taxas menores, quando comparado ao disponível no microbioma. Analisamos o microbioma de *S. leucops* em uma isolinhagem mantida por 12 anos em cultivo no laboratório, e variações em seu cultivo causado pela pandemia (cultivo em domicílio) e coleta da natureza atualmente. O microbioma de *S. leucops* não é rígido, como no caso de outro Catenulida que tem o microbioma bem estudado o *Paracatenula sp.* Esta tese de doutoramento, também traz o sequenciamento de DNA e RNA, além dos dados prévios de suas respectivas montagens e anotações. As sequências são provenientes das metodologias Illumina, Oxford Nanopore e IonTorrent S5.

Abstract

Description of the genome, transcriptome and microbiome of *Stenostomum leucops* (Platyhelminthes, Catenulide), focusing on phylogeny, regeneration, ecological plasticity and transposable elements.

AUTHOR: Marcos Trindade da Rosa

ADVISOR: Elgion Lucio da Silva Loreto

Platyhelminthes is the phylum with the highest adaptive radiation of invertebrates, and is currently divided into five classes: Catenulide, Rhabditophora, Monogenea, Cestoda and Trematoda. Phylogenetic analyses, using 12 mitochondrial genes from 165 species of Platyhelminthes, position Catenulide as a sister group to Rhabditophora. Long read sequencing revealed, in some Platyhelminthes species, that mitogenomes can be larger than those presented by short read sequencing. *S. leucops* has few differentiated tissues and is rich in stem cells, responsible for its regeneration and reproduction. The literature presents contradictory descriptions about regenerative aspects of *S. leucops* in the zooid formation, when the anterior cephalic part and the posterior body are removed. The experiments redone by us, found that all the regeneration ways described in the literature, can be obtained. The microbiome of *S. leucops* seems to go beyond simple food for the worm, as it has the capacity to express DNA contained in bacteria (plasmid with *gfp* marker). *S. leucops* is also able to utilize “free DNA” present in its environment, even at lower rates, when compared to that available in the microbiome. We analyzed the microbiome of *S. leucops* in an isolate maintained for 12 years in laboratory culture, and variations in its cultivation caused by the pandemic (home cultivation) and collection from nature today. The microbiome of *S. leucops* is not rigid, as in the case of another Catenulide that has a well-studied microbiome, *Paracatenula* sp. This doctoral thesis also brings DNA and RNA sequencing, in addition to previous data from their respective assemblies and annotations. The sequences come from the Illumina, Oxford Nanopore and IonTorrent S5 methodologies.

Lista de Abreviaturas

a.a- Aminoácidos

CDS(s)- Coding DNA sequence (sequência(s) de DNA codificante(s))

Covid- Corona virus disease (doença causada pelo vírus SARS-CoV-2)

DNA- Ácido Desoxirribonucleico

gfp- Green fluorescent protein (proteína fluorescente verde)

kb- kilo (quilo) bases * 1 000pb.

LNCC/MCT- Laboratório Nacional de Computação Científica/Ministério da Ciência e Tecnologia.

mRNA- Ácido ribonucleico mensageiro

pb- Pares de bases

RNA- Ácido ribonucleico

TE(s)- Elemento(s) de Transposição

tRNA- Ácido ribonucleico transportador

Sumário

1. Sumário.....	17
2. Capítulo 1 (Introdução Geral.).....	19
3. Objetivos.....	28
4. Capítulo 2 (A história evolutiva dos Platyhelminthes contada pelos mitogenomas.).....	29
5. Capítulo 3 (Revisiting the regeneration of <i>Stenostomum leucops</i> (Catenulida, Platyhelminthes.).....	54
6. Capítulo 4 (The Catenulida flatworm can express genes from its microbiome or from the DNA it ingests.).....	63
7. Capítulo 5 (<i>Stenostomum leucops</i> (Catenulida, Platyhelminthes) has a flexible microbiome in time and space.).....	75
8. Capítulo 6 (Análise do genoma e transcriptoma de <i>Stenostomum leucops</i> .).....	90

Capítulo 1
Introdução Geral

Platelmintos

O filo 'Platyhelminthes' “vermes-chatos” (Hyman 1951), foi inicialmente denominado de 'Plathelminthes' 1873, embora ambas as terminologias são aceitas (Ehlers e Sopott-Ehlers 1995). Os Platyhelminthes são acelomados, com ausência de órgãos circulatórios, excretores, e apresentam osmorregulação por protonefrídios (Marcus 1945). O filo é monofilético e está dividido atualmente em cinco classes (Figura 1): Catenulida, Rhabditophora, Monogenea, Cestoda e Trematoda (Schoch et al. 2020). Análises baseadas em marcadores moleculares e biologia do desenvolvimento, excluíram Acoela e Nemertodermida do filo Platyhelminthes (Telford et al. 2003, Wallberg et al. 2007, Durán e Egger 2012).

A classe Rhabditophora é composta de diversos grupos de organismos de vida livre, com capacidade de secretar muco para manter umidade e proteger seu epitélio. Rhabditophora apresenta cinco ordens mais estudadas e que possuem substanciais dados moleculares, nos bancos de dados, que permitem algumas análises evolutivas. São elas: Macrostomorpha, Tricladida, Polycladida, Rhabdocoelae e Bothrioplanida. A classe apresenta reprodução sexuada, com a presença de espermatozoides biflagelado, com exceção de Macrostomorpha (Morrow 2004). A biologia do desenvolvimento colocam Macrostomorpha e Polycladida como grupos basais dos Rhabditophora, por possuírem clivagem espiral e ovos entolecitais (Egger et al. 2015). Dados moleculares aproximam os Tricladida e Bothrioplanida dos Neodermatas (classes: Trematoda, Monogenea e Cestoda) (Egger et al. 2015).

Neodermata é uma superclasse monofilética, composto pelas classes representadas por organismos endoparasitas (Cestoda e Trematoda) e ectoparasitas (Monogênea) (Schockaert et al. 2008). A origem do primeiro grupo parasita de platelminto ocorreu aproximadamente no período Mesozóico/Paleozóico, antes mesmo do surgimento do grupo dos peixes modernos (Kearn 1994). O ancestral teve uma origem bentônica, de vida livre, mas com adaptação nas ferramentas de fixação, assim surgiu os primórdios dos ganchos para fixação. A fixação mais eficiente na forma de ganchos, ocorreu nos primeiros Monogêneas (Kearn 1994). O ectoparasitismo parece ser um caráter ancestral e o endoparasitismo derivado (Park et al. 2007). Os demais Neodermata (Cestoda e Trematoda) são endoparasitas, e com hospedeiros intermediários para seu desenvolvimento (Kmentová et al. 2016).

A classe menos estudada de Platyhelminthes é Catenulida, que traz mais controvérsias. O posicionamento filogenético, reprodução (preponderantemente assexuada), aspectos regenerativos, moleculares (genoma, transcriptoma, mitogenoma, interações microbiológicas ...), ainda carecem de dados e muitos inexistem. A proposta desta Tese é trazer alguns novos dados, visando ampliar o conhecimento sobre um membro da classe Catenulida, *Stenostomum leucops*.

Filogenética

O genoma mitocondrial (mitogenoma) dos metazoários é o marcador mais utilizado para reconstruções filogenéticas em vários níveis taxonômicos. Os mitogenomas têm características importantes para sua ampla utilização com confiabilidade em análises evolutivas. A obtenção do recurso (DNA mitocondrial) é facilitada pela abundância do material em cada célula, além de sua ortologia entre os genes codificantes, com taxas evolutivas diferentes aos marcadores nucleares, com baixa recombinação (Gissi et al. 2008; Elson e Lightowlers 2006). A informação das sequências nucleotídicas, os códons gerados, e a ordem dos genes mitocondriais, podem ser muito informativa sobre relações filogenéticas (Lavrov e Lang 2005). As mudanças na ordem gênica nos genomas mitocondriais não ocorrem com frequência, mesmo em longos períodos de tempo (Boore 1999). Os rearranjos nos mitogenomas também podem ter implicações na funcionalidade gênica e sua replicação. Sendo assim, esta informação pode ser muito importante para entendermos as relações filogenéticas (Gissi et al. 2008).

Os genomas mitocondriais de Platyhelminthes apresentam 12 genes codificantes, o que difere da maioria dos metazoários que, em geral, possuem 13 CDSs (do inglês coding DNA sequence). Isto ocorre pois os Neodermata perderam o gene *Atp8*, em seu mitogenoma durante a evolução (Gissi et al. 2008). No entanto, a CDS *Atp8* tem sido encontrada nos mitogenomas de Tricladida (Ross et al. 2016) e Catenulida (Egger et al. 2017, Rosa et al. 2017). O gene *Atp8* pode não ser localizado, muitas vezes, devido as técnicas convencionais de busca (similaridade e se tratar de um fragmento curto 100-200 pb). Além disso, Monnens et al. (2020) sugerem que alguns mitogenomas podem estar descritos e anotados incompletos por terem sido gerados por metodologias de sequenciamento com *reads* “curtos”, e novos sequenciamento utilizando *long reads* mostram genomas mitocondriais maiores e mais completos.

O posicionamento filogenético de Catenulida é discutido em vários ramos da biologia dos Platyhelminthes, atualmente os dados da biologia do desenvolvimento e molecular, apontam Rhabditophora e Catenulida como grupos irmãos (Peterson e Eernisse 2001; Jondelius et al. 2002; Telford et al. 2003, Larsson e Jondelius 2008). Catenulida era apontado como basal quando Acoela e Nemertodermatida ainda eram incluídos nos Platyhelminthes (Ehlers 1985). Catenulida também foi apresentado como grupo irmão dos Bilateria, em vez de grupo irmão de Rhabditophora (anteriormente Rhabditophora + Neodermatas), junto a Acoela e Nemertodermida (Haszprunar 1996). Catenulida também já foi parafilético com os Platyhelminthes, e basal aos demais Bilateria (Carranza et al. 1997). Estudos mais recentes, utilizando dados moleculares (COI, 18S, mitogenomas, RNAs) trouxeram Catenulida + Rhabditophora como grupos irmãos dentro de

Platyhelminthes, com a remoção de Acoela e Nemertodermatida dos Platyhelminthes (Carranza et al. 1997, Litvaitis e Rohde 1999, Telford et al. 2003, Wallberg et al. 2007 Egger et al. 2015, Laumer et al. 2015).

A maior parte dos trabalhos filogenéticos de Platyhelminthes usam poucos marcadores moleculares (COI, 18S e 28S), ou mitogenomas dos grupos específicos (classe, família, gêneros). A utilização de poucos marcadores, e/ou mitogenomas, trazem parafilias dentre famílias de Rhabdtophora, Cestoda e Trematoda (apresentadas e discutidas no capítulo 2). A classe Monogênea também parece ser parafilética em suas duas subclasses Monopisthocotylea e Polyopisthocotylea (Justine 1998, Littlewood et al. 1999, Laumer e Giribet 2014). No presente momento há centenas de mitogenomas de Platyhelminthes disponibilizados no GenBank, embora o maior número de espécies representantes são dos Neodermata (parasitas de importância médica e econômica). Os grupos que evolutivamente ramificaram primeiro, que contemplam os organismos de vida livre (Rhabdtophora + Catenulida) ainda carecem de mais dados.

Os mitogenomas de Catenulida apresentam peculiaridades em relação aos demais Platyhelminthes, como a presença de CDSs em ambas as fitas, e o tRNA lisina (tRNA-K) que apresenta característica de plesiomorfismo com os demais invertebrados não vermiformes. Os invertebrados não platelmintos tem como códons AAA lisina (Lys), já nos platyhelminthes este aminoácido é codificado apenas por CTT, e AAA codifica aspargina (Asn) (Telford et al. 2000, Rosa et al. 2017; Egger et al. 2017). Os mitogenomas descritos de Rhabdtophora apresentam codificação na mesma fita, ou seja, transcritos na mesma direção, característica incomum presente em apenas cinco filos (Annelida, Brachiopoda, Platyhelminthes ‘não Catenulidas’, Cnidaria e Porifera) (Gissi et al. 2008). Uma contribuição desta tese, presente no capítulo 2, é uma análise filogenética ampla dos platelmintos, com o mitogenoma de 165 espécies com a inclusão dos mitogenomas de Catenulida: *S. leucops* (Rosa et al. 2017) e *S. sthenum* (Egger et al. 2017)

O trabalho trazido no capítulo 2 apresenta o arranjo gênico de Catenulida/Stenostomidea, que também ficou mais claro com o sequenciamento Nanopore (*long reads*), e traz a possível ordem gênica em nível de família. Mitogenomas de *long reads* de *Schistosoma* apresentam repetições de regiões (CDSs, tRNAs, regiões não codificantes), que podem atrapalhar na montagem, descrição, estudos evolutivos, e ausência dados mais completos (Kinkar et al. 2021). Neste trabalho ampliamos os dados de sequenciamento de *S. leucops*, também realizando sequenciamento *long reads*, e através dessa metodologia também podemos mostrar uma extensão da estrutura do mitogenoma de *S. leucops*. A reconstrução filogenética mais completa, permite a visão mais clara da história evolutiva mitocondrial dos Platyhelminthes e posicionamento do ramo Catenulida.

Catenulida

A classe Catenulida (Meixner 1924) é composta por organismos bentônicos de vida livre, que habitam água doce, com poucos milímetros de tamanho (0,2-3 mm) (Larsson et al. 2008). Os caracteres morfológicos de Catenulida para fins diagnósticos são escassos, devido à presença de poucos tecidos diferenciados (Larsson e Jondelius 2008). A reprodução é preponderantemente assexuada, com espermatozoide aflagelado, mas é descrito que estes organismos são capazes de desenvolver casulo reprodutivo (Borkotta 1970). Os trabalhos mais atuais de Catenulida trazem eles como assexuados, mas sendo controverso em Borkotta (1970). O desenvolvimento é direto, sem ânus, com cérebro e abertura bucal na parte anterior do corpo. Os estatocistos de Catenulida e Rhabditophora, divergem em sua morfologia, com evoluções provavelmente independente em sua estrutura (Ehlers 1991).

Os organismos mais estudados dentre os Catenulida são dos gêneros *Catenula*, *Paracatenula* e *Stenostomum*. Atualmente, há pouco mais de 100 espécies descritas (Larsson e Jondelius 2008).

“*Stenostomum leucops*”

Stenostomum leucops Dugès 1828, é uma espécie de vida livre, cosmopolita, com 0,5-2mm de comprimento e reprodução preponderantemente assexuada (Noreña et al. 2005). *S. leucops* apresenta alta taxa reprodutiva, apresentando a formação completa dos zooides em aproximadamente 48 horas quando mantidos a 28°C. São facilmente cultivados em laboratório, tendo potencial como organismo experimental (Rosa et al. 2015). O resultado desta reprodução são gerações isogênicas em curto prazo de tempo (Hyman 1951). A classificação morfológica em nível de espécie pode ser difícil, devido aos poucos caracteres morfológicos diagnósticos. Outro caracter morfológico é a paratomia (quanto ao número de zooides seguido de fissão), e os gânglios faríngicos (forma de bastonetes ou ovalados, distribuídos pela sua extensão) (Noreña et al. 2005).

A paratomia ocorre pelo crescimento corpóreo em sentido axial e formação cefálica em partes medianas do corpo (zooides), as contrações corpóreas auxiliam na fissão, separando a região posterior do(s) zooide(s) anterior(es). A cadeia de zooide formada em *S. leucops* SL0-sm01 é modificada pelo microbioma associado (2-5+ zooides) (Rosa et al. 2015), que difere do descritos como característica da espécie, que seriam de formação de apenas dois zooides (Noreña et al. 2005). A paratomia é um processo similar a metamerização, e pode ter bases genéticas comuns à segmentação, que são observadas em grupos mais derivados dos Platyhelminthes (Neodermatas). Um vídeo demonstrando o processo de formação de zooides e o processo de paratomia pode ser acessado em <https://www.youtube.com/watch?v=zzE6VM0Uw6I>.

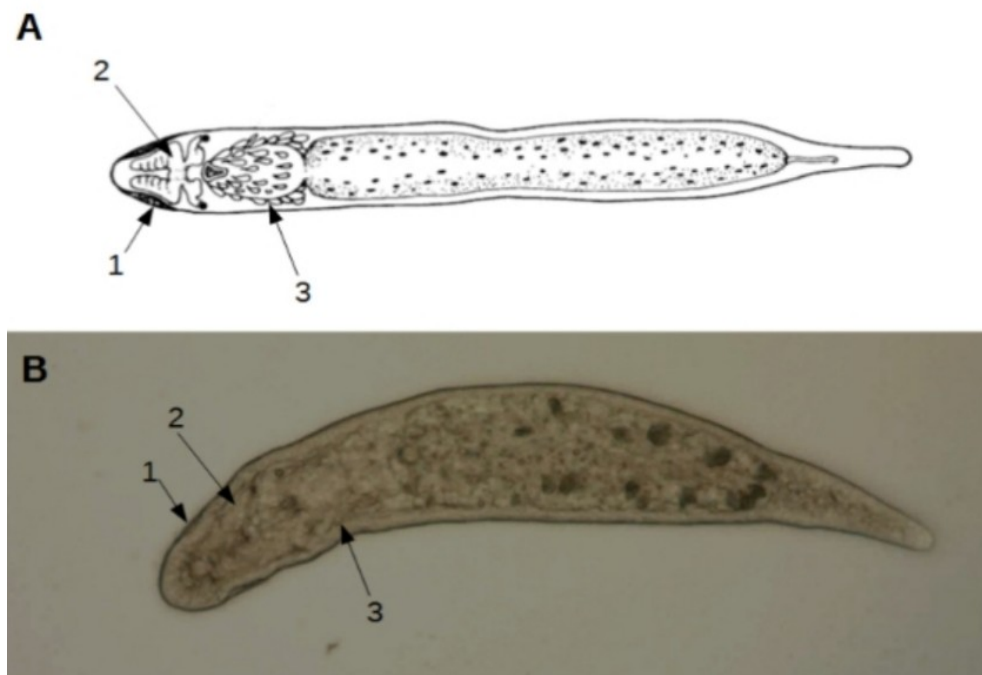


Figura 1: A) Representação esquemática de estruturas morfológicas de *S. leucops* segundo Noreña (2005), B) Fotografia Rosa et al. (2015) de *S. leucops*; 1: fossos ciliares, 2: gânglios cerebrais e 3: glândulas da faringe.

Regeneração

Platyhelminthes tem sido usados em ensaios de regeneração há mais de um século (Fredrick et al. 2018). A regeneração é outra característica que torna *S. leucops* potencial organismo experimental, por possuir uma alta taxa regenerativa (Van Cleave 1929). A regeneração ocorre devido à presença de células-tronco (*stem cells*), denominadas comumente de neoblastos. Entretanto, os padrões regenerativos desta espécie apresenta resultados contraditórios.

As divergências na regeneração durante a formação de zóide/paratomia foram trazidas por três estudos. Os estudos removiam a região cefálica anterior e região corpórea posterior. Ritter e Congdon (1900) descrevem que, quando um organismo em processo de formação de zóide é seccionado removendo sua região cefálica anterior, a parte cefálica posterior pode migrar pelo corpo até assumir a posição anterior. Child (1903) descreve que em um verme ao ser seccionado há formação de uma região cefálica anterior e degeneração da posterior. Nuttycombe e Waters (1938) apontam que ocorre uma nova formação cefálica anterior e formação de uma região corporal após a região cefálica posterior. As divergências podem ser associada a pesquisas com diferentes espécies, mas todas sendo nomeadas de *S. leucops*, devido aos poucos caracteres morfológicos diagnósticos. Nuttycombe e Waters (1938) apontam que *S. leucops* não seria uma espécie válida devido à divergência de resultados. Posteriormente, estudos feitos por caracteres morfológicos detalhados de Borkott (1970) e Noreña et al (2005) validaram a espécie. Dados e estudos moleculares recentes

utilizados por Larsson et al. (2008) e Rosa et al. (2015) apontam que *S. leucops* pode se tratar de um complexo de espécies.

O estudo apresentado no capítulo 3, traz os aspectos regenerativos de *S. leucops*. Todos os processos regenerativos descritos anteriormente são encontrados na cepa SL0-sm02. Outro fator que influencia a regeneração em Platyhelminthes é a microbiota envolvida, pois seus metabolitos podem agredir os neoblastos e inibir o processo (Fredrick et al. 2018). Ou, diferentemente, uma microbiota apropriada pode induzir a formação do maior número de zooides (Rosa et al. 2015). *S. leucops* depende da proliferação de seus neoblastos para a regeneração e reprodução, modificações na microbiota podem afetar diretamente estes aspectos.

Microbioma

As características externas e internas de Platyhelminthes criam ambientes únicos para comunidades bacterianas, pois secretam quantidades significativas de muco em suas superfícies externas, e secreções internas exteriorizadas pela faringe (muco digestivo → auxiliar pré-digestão externa) (Hayes 2017, Fredrick et al. 2018). Os Catenulida também utilizam a microbiota como fonte de alimento, ou até como endossimbiontes no caso de Paracatenula (Gruber-Vodicka et al. 2011).

O estudo da microbiota de *S. leucops* é fundamental para entendermos sua biologia. Os aspectos fenotípicos de *S. leucops* podem ser influenciados pelo microbioma, como por exemplo, a formação dos zooides (Rosa et al. 2015). Uma pergunta que estamos propondo responder é se há uma ligação estrita entre *S. leucops* e seu microbioma. A pandemia causada pelo COVID-19 proporcionou que a linhagem de *S. leucops* mantida em nosso laboratório fosse mantida também em cultura doméstica (em casa) com uso de água de torneira para facilitar a manutenção, mas seguindo o mesmo método que no laboratório (com exceção da água). Os aspectos biológicos como a formação na cadeia de zooides tiveram alteração. Aproveitamos estas diferenças de cultivo para fazer um estudo do microbioma, comparando as linhagens de laboratório, mantidas em cultivo doméstico e uma população retirada da natureza (capítulo 5).

Genoma

O sequenciamento e análise de genomas produz um conjunto de dados muito importante para entendermos a biologia e a evolução dos organismos. Os genomas de Platyhelminthes têm grande atratividade para diversos estudos, visto que parte dos organismos deste grupo tem importância médica, como parasitas, ou pela presença de células totipotentes para estudos sobre regeneração (Hahn et al. 2014), além de serem o grupo de invertebrados com uma das maiores variedades de

adaptações (Fontenla et al. 2021). Um exemplo é a planária *Schmidtea mediterranea*, que tem um banco de dados exclusivo, Genomic Database (SmedGD) criado em 2007 (Robb et al. 2008). Este organismo é a principal referência para estudos sobre regeneração dentre os Platyhelminthes.

Alguns estudos atuais começaram a explorar uma diversidade maior de Platyheminthes com capacidades regenerativas, por exemplo, *Macrostomum lignano* que é capaz de formar partes perdidas em três dias (Egger et al. 2006). Nos dados moleculares (genoma e o transcriptoma) de *M. lignano* não foram encontrados sinais dos principais fatores de pluripotência dos mamíferos (Oct4 / Pou5f, Klf4 e c-Myc) (Wasik et al. 2015). Os fatores de pluripotência também não foram encontrados em *S. leucops* em nossas análises prévias. O entendimento da capacidade regenerativa nos Platyhelminthes e evolução dos mesmos, devem avançar com dados moleculares de vários modelos (Wasik et al. 2015). As análises comparativas entre os genomas/transcriptomas de *S. mediterranea*, *M. lignano* e *S. leucops*, podem ser importantes ferramentas para elucidar alguns fatores regenerativos, desenvolvimento, evolutivos, entre outros. A evolução da metamerização/segmentação desenvolvida primeiramente em Cestoda, pode ser melhor compreendida, com buscas de regiões homologas com organismos evolutivamente mais basais.

O capítulo 6 apresenta uma visão preliminar dos passos que foram e ainda serão tomados, para as três metodologias de sequenciamento feitas Illumina (*reads* curtos), IonTorrent S5 (transcritos-RNA) e Oxford Nanopore (*long reads*). As montagens genômicas e transcriptômicas similares, ou superiores, quando comparadas a composição CDSs (sequências de DNA codificantes), em relação aos já disponibilizados para os organismos filogeneticamente próximos (*Macrostomum lignano* e *Schistosoma mansoni*). O tamanho do genoma estimado pela montagem de sequenciamento *Nanopore*, apresenta ser próxima ao que se estipula, em peso molecular, para a família Stenostomidae. Análises, através das três metodologias de sequenciamento, têm potencial para gerar dados sobre organização genômica e aspectos evolutivos para outros artigos em um futuro momento, mas ainda carecem de uma descrição mais cautelosa.

Elementos de transposição (TEs)

Elementos de transposição (TEs) são sequências de DNA com capacidade de mudar sua localização dentro do genoma. Quando os TEs possuem informação genética para produção da maquinaria enzimática para sua própria transposição, são denominados autônomos. Quando utilizam a maquinaria de outro elemento são denominados não-autônomos (Wicker et al, 2007). Os TEs foram descobertos por McClintock em 1950, no milho, por causarem alterações no padrão de coloração dos grãos. Desde então, muitos estudos mostram que os genomas são dinâmicos, e que a

transposição destas sequências de DNAs, tem grande influência nos aspectos biológicos e evolutivos dos organismos hospedeiros (Dolgin e Charlesworth 2006, Varani et al. 2015).

Finnegan (1989) propôs um sistema de classificação para os TEs dividindo-os em Classe I, ou retrotransposons, que seriam elementos que se movem por um intermediário de RNA, usando uma enzima transcriptase reversa, homóloga as usadas por retrovírus. Estes elementos são transcritos para RNA e copiado reversamente para DNA, sendo posteriormente inserido em uma nova região do genoma. Já os elementos de Classe II, também conhecidos como transposons, agrupam elementos que se transpõem usando um intermediário de DNA, normalmente usando uma enzima chamada transposase que “corta” o elemento da sua posição inicial e o “insere” em uma nova posição do genoma. Embora os helitrons e elementos relacionados, são elementos de classe II e não utilizem o mecanismo “corta e cola” para transposição. Com os avanços das técnicas moleculares e de sequenciamento, a diversidade de TEs descritos vem aumentando, como por exemplo, o banco de dados de transposons no Repbase já apresenta 56.000 sequências, subdivididas em 65 superfamílias (Bao et al. 2015). Com o aumento da diversidade de sequências descritas, novos padrões classificatórios também estão sendo propostos, com mais subdivisões. Alguns pesquisadores da área sugeriram a criação de um comitê para organizar, ajudar a padronizar e melhorar estas classificações (Arensburger et al. 2016).

Os genomas de Platyhelminthes de vida livre, descritos até o momento, são ricos em sequências repetitivas e elementos de transposição. Por exemplo, em *M. lignano* apresenta 75% do genoma composto por sequências repetitivas, como regiões de microssatélites e elementos de transposição (Wasik et al. 2015). Mesmo que até o momento não se tenha estudos suficientes sobre o material genético transcricional, gerados destas regiões repetitivas, os autores levantam a hipótese que elementos de transposição se “aproveitam” do sistema de silenciamento destes Platyhelminthes. Nesses Platyhelminthes o sistema de silenciamento (rotas Piwi, Argonautas, micro-RNAs) tem como principal função a manutenção dos neoblastos, atuando na regeneração e crescimento em *Macrostomum lignano* (Palakodeti et al. 2008, Zhou et al. 2015). Assim, abrindo falhas no silenciamento de elementos transponíveis, permitindo a sua proliferação nestes genomas. A atividade dos TEs podem ajudar a promover muitos efeitos em células somáticas: positivo (promover plasticidade, diversidade em organismos assexuados), efeitos neutros (sem interferência benéfica e nem prejudicial ao funcionamento celular), ou negativo (fazendo com que a célula entre em apoptose ou câncer) (revisão em Loreto e Pereira 2017).

No capítulo 6 apresentamos uma análise preliminar dos genomas e transcriptoma de *S. leucops*. As análises precisam ser refinadas, uma vez que devido a pandemia Covid-19, houve dificuldades de obtenção das amostras e para o próprio sequenciamento *long reads* (Nanopore).

Objetivos gerais

-Ampliar os conhecimentos da biologia evolutiva, do desenvolvimento, ecologia e genômica de *S. leucops*, por ser basal evolutivamente, é um organismo promissor para estudos comparativos com outros Platyhelminthes.

Objetivos específicos

-Filogenética mitocondrial de Platyhelminthes.

-Capacidade regenerativa de *S. leucops*, com relação aos resultados contraditórios descritos na literatura para este processo.

-Averiguar a interação *S. leucops* com bactérias do gênero *Bacillus* e *Escherichia*, que em abordagens anteriores em nosso laboratório indicavam que este verme tem a capacidade incomum de expressar genes contidos em DNA presente em seu meio de cultura ou no seu microbioma.

-Descrição da sucessão microbiológica das comunidades bacterianas no espaço e tempo, em cultivo de *S. leucops* em laboratório e comparar com populações obtidas na natureza.

-Apresentar os dados já analisados referentes ao genoma e transcriptoma de *S. leucops*.

Capítulo 2

A história evolutiva dos Platyhelminthes contada pelos mitogenomas.

(Artigo em fase de redação)

A história evolutiva dos Platyhelminthes contada pelos mitogenomas.

Resumo

O filo Platyhelminthes atualmente, está dividido em cinco classes: Catenulida, Rhabditophora, Monogenea, Cestoda e Trematoda, que apresentam os mais diversos estilos de vida, podendo ser considerado um modelo de radiação adaptativa. Com vida livre temos Catenulida, com seus representantes de poucos milímetros, assexuados límicos; Rhabditophora que habitam a água doce, salgada e ambiente terrestre. Platyhelminthes também é composto por parasitas os Neodermatas, que vão de parasitas externos com os Monogenea, aos Cestoda e Trematoda como parasitas internos. As relações filogenéticas deste grupo segue em discussão. Os mitogenomas de metazoários podem ser utilizado para reconstruções filogenéticas, possuindo confiabilidade em características evolutivas, facilidade de obtenção, ortologia entre os genes codificantes, evolução em taxas diferentes aos marcadores nucleares, além da baixa recombinação. Através das 12 CDSs dos mitogenomas de Platyhelminthes, das 165 espécies depositados no NCBI, reconstruímos a história evolutiva mitocondrial deste filo. A filogenética mitocondrial posicionou Catenulida como grupo irmão de Rhabditophora. Monogenea aparece como parafilético em suas subclasses Monopisthocotylea e Polyopisthocotylea. Polyopisthocotylea recuperou como grupo irmão Cestoda e Trematoda. Todos as classes apresentaram parafilias em suas famílias pela análise de sequências de blocos conservado de aminoácidos e, além disso, convergem pela ordem de disposição dos genes dentro do mitogenoma. O sequenciamento Nanopore de *S. leucops* apresentou uma expansão no tamanho mitogenoma, e apresenta duplicações quando comparado aos mitogenomas montados e anotados disponíveis no NCBI, como já observado para família *Schistosomatidae*. Este sequenciamento possibilitou estabelecer a estruturação comparativa na família *Stenostomidae*.

Palavras chaves: Platyhelminthes, mitogenomas, reconstrução filogenética, ordem genica e Nanopore.

Introdução

Os Platyhelminthes são animais metazoários com composição corpórea simples, conhecidos como vermes de corpo achatado (Hyman 1951). Estes vermes estão distribuídos por quase todos ecossistemas da terra e, também, são modelos de radiações adaptativas (Park et al. 2007). Os estudos empregando marcadores moleculares e aspectos relacionadas com a biologia do desenvolvimento apontaram para a monofilia dos Platyhelminthes, e retiraram Acoela e Nemertodermida como membros desse filo (Jondelius et al. 2002, Telford et al. 2003, Wallberg et

al. 2007, Larsson et al. 2008, Egger et al. 2015, Laumer et al. 2015). Os Platyhelminthes estão divididos em cinco classes: Catenulida, Rhabditophora, Monogenea, Cestoda e Trematoda (Schoch et al. 2020).

Os Catenulida são organismos de vida livre, com 0,2 a 2 mm de comprimento, possuem poucos caracteres morfológicos diagnósticos (Norenã et al. 2005), os espermatozoides são aflagelados (Sopot-Ehlers e Ehlers, 1999), e reprodução preponderantemente assexuada por formação de zooides (Larson et al 2008). Já em Rhabditophora a espermatogênese, em quase todas as famílias, gera espermatozoides biflagelados, com exceção de Macrostromida, que também apresenta ausência de flagelos (Morrow 2004, Willems et al. 2009). Catenulida, em grande parte dos estudos, aparece como uma ramificação basal ou grupo irmão de Rhabditophora (Jondelius et al. 2002, Telford et al. 2003, Wallberg et al. 2007, Larsson e Jondelius 2008, Egger et al. 2015, Laumer et al. 2015).

A classe Rhabditophora é composta de diversos grupos de organismos de vida livre, como os Macrostromorpha, Tricladida, Polycladida, Rhabdocoela e Bothrioplanida. Características associadas à biologia do desenvolvimento colocam Macrostromorpha e Polycladida como grupos basais dos Rhabditophora, pois ambos possuem clivagem espiral e ovos entolecitais, sendo estes prováveis caracteres primitivos (Egger et al. 2015). Os Tricladida têm forte apoio para uma posição mais próxima de Neodermata e Bothrioplanida, pois nenhum Tricladida mostra qualquer sinal de clivagem espiral em sua embriogênese inicial. Dados moleculares também aproximam os Tricladida e Bothrioplanida dos Neodermatas (Monogenea, Cestoda e Trematoda) (Egger et al. 2015).

Os clados mais derivados são formados por uma vasta gama de vermes parasitas: Monogenea, Cestoda e Trematoda (Neodermatas). A evolução do primeiro grupo de Platyhelminthes parasitas se deu próximo ou antes mesmo do aparecimento dos peixes modernos, através de um ancestral Platyhelminthes bentônico, com modificações no mecanismo de adesão ao substrato. A capacidade de adesão ao substrato é comum em Rhabditophora, com a adaptação e evolução das ferramentas de fixação para ganchos, provavelmente dando origem ao primeiro grupo de Platyhelminthes parasitas, Monogenea (Kearn 1994). Análise filogenética empregando genes mitocondriais apontam que Cestoda é grupo irmão de Trematoda e divergiram de um ancestral comum com os Monogenea. Os Monogenia apresentam hábito de ectoparasitismo, com hospedeiro único, sem intermediários. Os demais Neodermata são endoparasitas e com intermediários para seu desenvolvimento (Kmentová et al. 2016). Assim, o ectoparasitismo parece ser um carácter ancestral e o endoparasitismo derivado (Park et al. 2007). Trabalhos descritivos de genoma mitocondrial (mitogenoma) de espécies de Monogenea, com reconstruções filogenéticas pontuais de classes, apresentaram parafilia das subclasses Monopisthocotylea e Polyopisthocotylea (Justine 1998, Littlewood et al. 1999, Laumer e

Giribet 2014). Outro aspecto digno de nota é que análises moleculares apontam parafilias entre famílias que compõe Monogenea (Mollaret et al. 2000; Zhang et al. 2018), Cestoda (Xi et al. 2018) e Trematoda (Locke et al. 2018).

O genoma mitocondrial (mitogenoma) dos metazoários é o marcador mais utilizado para reconstruções filogenéticas em vários níveis taxonômicos. Podemos citar algumas características desse genoma como razões para sua ampla utilização: a confiabilidade, características evolutivas, facilidade de obtenção (abundância de material em cada célula), ortologia entre os genes codificantes, evolução em taxas diferentes aos marcadores nucleares, além da baixa recombinação (Gissi et al. 2008; Elson e Lightowlers 2006). A informação associada a sequência nucleotídica, os códons empregados, e a ordem dos genes mitocondriais pode ser muito informativa sobre relações filogenéticas, por apresentarem baixa divergência entre os clados (Lavrov e Lang 2005). As mudanças na ordem gênica no genoma mitocondrial não ocorrem com frequência, mesmo em longos períodos de tempo (Boore 1999). Os rearranjos nos mitogenomas também podem ter implicações na funcionalidade gênica e na sua replicação. Sendo assim, estas informações podem ser muito importantes para entendermos as relações filogenéticas (Gissi et al. 2008).

Os genomas mitocondriais de Platyhelminthes apresentam 12 genes codificantes (CDS(s) do inglês, *coding DNA sequence*), diferindo da maioria dos mitogenomas dos metazoários que, em geral, possuem 13 CDSs. Isto ocorre pois os Neodermata perderam o gene *Atp8* durante a evolução (Gissi et al. 2008), e esta CDS não é encontrada em grande parte dos mitogenomas de Platyhelminthes depositados nos bancos de sequência. Trabalhos recentes, envolvendo Catenulida e Tricladida, apontam que os organismos destes grupos apresentam *Atp8*, mas que, muitas vezes pelas técnicas convencionais de busca por similaridade e, por se tratar de um fragmento curto (100-200 pb), este gene não é anotado (Monnens et al. 2020). Genomas mitocondriais de Platyhelminthes apresentam, em geral, de 13 a 17 kpb. Entretanto, Kinkar et al. (2021), estudando *Schistosoma*, através da metodologia de sequenciamento *long reads*, sugerem que os mitogenomas podem ser maiores, tendo regiões duplicadas, mas que montagens a partir de sequenciamento de *short reads* geram problemas de montagem, resultando montagens de genomas menores do que realmente são.

No momento, existem centenas de mitogenomas de Platyhelminthes disponíveis no GenBank, embora a maioria das espécies sejam representantes de Neodermata e poucos representantes de grupos basais. Catenulida, por exemplo, possui apenas dois mitogenomas completos e anotados. O objetivo deste trabalho foi pesquisar mitogenomas de Platyhelminthes disponíveis publicamente para traçar a história evolutiva dos genomas mitocondriais deste grupo, para a qual foi feita a reconstruções filogenéticas a partir de 165 mitogenomas Platyhelminthes depositados no GenBank

(NCBI). Além da filogenômica baseada em sequências gênicas mitocondriais, também analisamos a ordem e estrutura gênica dos tRNAs como caracteres importantes para o entendimento das relações evolutivas dos grupos dos Platyhelminthes. Também apresentamos uma montagem do mitogenoma de *S. leucops* obtido por sequenciamento *long reads*.

MATERIAIS E MÉTODOS

Sequenciamento de *long reads* pela metodologia Oxford Nanopore

A extração do DNA genômico partiu de 30.000 organismos *S. leucops*, e implementado conforme o protocolo fenol/clorofórmio (Sassi et al. 2005). A concentração final de 94 ng/μl de DNA, em 59 μl de água para construção das bibliotecas. As razões $260/280 = 1.84$ e $260/230 = 1.83$ pelo Nanodrop. As amostras foram secas em vácuo, na temperatura ambiente em tubos GenTegra™. O sequenciamento foi realizado em plataforma PromethION-Oxford Nanopore pela empresa GenOne, RJ, Brasil.

Filogenética mitocondrial

As análises filogenéticas foram realizadas a partir de duas matrizes. Primeiramente uma com 165 espécies de Platyhelminthes e os blocos conservados de 12 proteínas mitocondriais. Esta matriz foi construída pelos alinhamentos individuais de cada uma das 12 proteínas, através do ClustalW do pacote MegaX (Kumar et al. 2018). Os alinhamentos individuais passaram no software Gblock (Talavera e Castresana et al. 2007), para gerar blocos conservados informativos. As 12 matrizes foram concatenadas (*Nad5, Nad3, Atp6, Nad6, Nad1, Nad2, Co2, Co3, Cytb, Nad4, Cox1, Nad4l*), para gerar uma grande matriz com 1903 caracteres para cada organismo (material suplementar Alinhamento 1.nexus).

A segunda matriz foi feita com 166 espécies de Platyhelminthes, com seis das 12 sequências de aminoácidos das proteínas mitocondriais (*Nad1, Co2, Cytb, Nad5, Cox1, Cox3*), pois somente estes foram recuperados do RNAseq de *Microstomum lineare* (SRR1980039). A matriz contém 496 caracteres de aminoácidos (material supl. Alinhamento2.nexus), e seguem a mesma metodologia anteriormente descrita para a análise com os 12 genes.

As duas matrizes anteriormente descritas, foram usadas para reconstrução filogenética por Máxima Verossimilhança e análise Bayesiana. Para a reconstrução filogenética de Máxima Verossimilhança, empregamos o pacote MegaX. O modelo evolutivo mais apropriado para os dados foi calculado pelo mesmo pacote, seguido de 10000 réplicas. Para a reconstrução filogenética por análise Bayesiana utilizando o pacote Mrs. Bayes (Huelsenbeck et al. 2001), onde modelo evolutivo

empregado foi estipulado empregando o software ModelTest-NG (Darriba et al. 2019), através da ferramenta ModelProt teste, seguido de 3000000 réplicas.

Anotação dos mitogenomas

As sequências dos mitogenomas de 165 espécies de Platyhelminthes e suas respectivas anotações das 12 CDSs foram obtidas no NCBI (<https://www.ncbi.nlm.nih.gov/>) (Tabela supl. 1.2). Como grupo externo foi utilizado o rotífero *Brachionus plicatilis*. Para a segunda matriz com 166 Platyhelminthes e seis CDSs, a sequência de aminoácidos foi obtida através da montagem dos *reads* do transcriptoma de *Microstomum lineare* (SRR1980039). Neste caso, os *contigs* foram montados através do software Trinity (Grabherr et al. 2011). Os genes mitocondriais foram separados por uma busca Blast (material supl. Tabela1; Aba 1).

Ordem gênica

Os mitogenomas das 165 espécies de Platyhelminthes foram tabelados, e agrupados pela classificação filogenética a qual estão dispostos no NCBI (Tabela suplementar 1). A tabela apresenta o arranjo dos 12 genes na sequência em como eles estão dispostos em cada mitogenoma. Os arranjos foram ordenados partido de um gene selecionado aleatoriamente (*Cox1*), uma vez que os mitogenomas são circulares. Esta matriz foi utilizada para a análise da ordem gênica nos mitogenomas dos Platyhelminthes.

A ordem gênica do mitogenoma de *Stenostomum leucops* obtida a partir de sequenciamento *longs reads*, foi gerada através de um consenso de sete *reads*. Os *reads* foram selecionados pela alta similaridade com o mitogenoma já depositado (97-98% de identidade com regiões CDSs). Os sete *reads* foram anotados no Mitos Web Server (Donath et al. 2019), através dessas anotações feito o consenso estrutural.

tRNAs

Para análises dos tRNAs foram selecionados representantes de todos os diferentes ramos filogenéticos dos Platyhelminthes. As estruturas secundárias dos tRNAs foram geradas empregando o software Mitos Web Server (Donath et al. 2019). As estruturas secundárias e região do anticódon foram utilizadas para análises comparativas.

Os Catenulida apresentam somente dois mitogenomas até a data de 11/11/2020 (*S. leucops* e *S. sthenum*), em uma busca usando a ferramenta Blast, também foi encontrado fragmentos mitocondriais nos transcritos de RNA-seq de *S. sp* SM09 (SRR9915028), que possibilitou isolar o tRNA^k-lisina, para participar das análises.

Resultados

Mitofilogenômica de Platyhelminthes

As análises filogenéticas implementadas por Máxima Verossimilhança e, também por análise Bayesiana, contendo as 165 espécies (12 CDSs), recuperou um ramo monofilético com todos os Platyhelminthes (Figura 1.2:A). As árvores filogenéticas completas podem ser encontradas nas Figuras Suplementares 1.2 e 2.2. Os Platyhelminthes se separaram em dois grandes ramos, um com organismos de vida livre (Catenulida + Rhabditophora) e outro com os parasitas Neodermatas (Monogenea, Cestoda e Trematoda). O mesmo pode ser observado para a matriz com 166 espécies (seis CDSs) (Figura 1: B). O ramo da filogenia com os 12 blocos conservados, agrupou *Macrostomum lignano* (Rhabditophora) com os Catenulida (Figura 1.2: A). Já quando utilizamos os seis blocos das 166 espécies, recuperamos a monofilia de Rhabditophora, e este apresenta Catenulida como grupo irmão (Figura 1.2: B). Já no ramo dos Neodermata (Monogenea, Cestoda e Trematoda), em ambas filogenias, Monogenea aparece como parafilético em suas subclasses Monopisthocotylea e Polyopisthocotylea. Polyopisthocotylea aparece como grupo irmão aos demais Neodermatas. Monopisthocotylea se ramifica e agrupa como irmão a Cestoda + Trematoda. As modificações nas ordens gênicas foram transpostas e concordantes com a formação das monofilias e parafilias, estes resultados serão abordadas mais detalhadamente em cada clado e seus subtópicos.

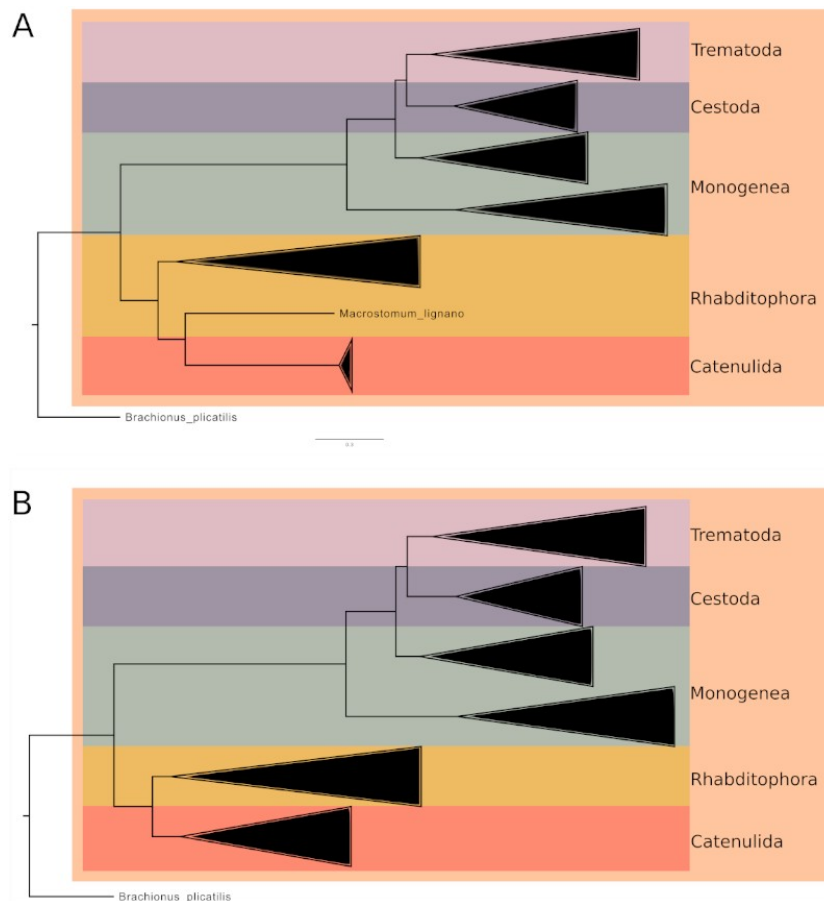


Figura 1.2: A) Filogenia colapsada (baseada na figura suplementar 1) em que se pode observar as em suas classes, gerada através de 12 CDSs de 165 mitogenomas Platyhelminthes + *Brachionus plicatilis*. B) Filogenia colapsada por suas classes, gerada através de seis CDSs de 166 mitogenomas Platyhelminthes + *Brachionus plicatilis*.

Ordem gênica no gênero *Stenostomum*

A ordem gênica encontrada nos mitogenomas no gênero *Stenostomum* obtida para dois representantes que possuem dados no NCBI, *Stenostomum leucops* e *Stenostomum sthenum* encontra-se na Figura 2. O gênero *Stenostomum* é o único representante de Catenulida com mitogenomas completos/semi-completos depositados. A ordem gênica de *S. leucops* e *S. sthenum* apresentam três blocos conservados. Dois blocos mantêm todos genes em sintonia de ordem e posição, um bloco conservado menor (*Cox1*/*Nad3*) e o bloco conservado maior (*-Cox2*/*-Nad4L*/*Nad4*/*Atp6*/*Nad5*/*Nad2*). O terceiro bloco do mitogenoma apresenta uma inversão na ordem e sentido das regiões CDSs.

Sequenciamento por *long reads* do genoma de *S. leucops* permitiu verificar uma expansão do tamanho do mitogenoma em mais 26.775 pb. A ordem gênica do mitogenoma e sua estruturação

também ficou mais clara. A ordem das CDSs apresentou semelhança com *S. sthenum*, e inserção de repetição de regiões CDS, com ordem semelhantes, porém com a duplicação de 5 CDSs (Figura 2).

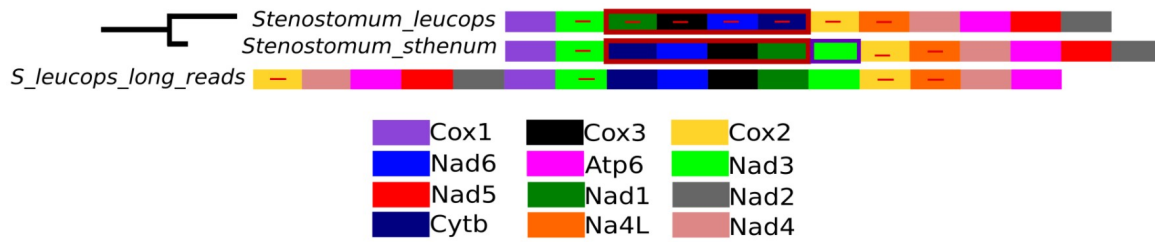


Figura 2.2: Ordem gênica observada no mitogenoma em *Stenostomum leucops* e *S. sthenum*. Bloco vermelho: inversão do sentido das CDSs, bloco contornado azul: Inserção de bloco CDSs, traço vermelho: transcrição em sentido contrário.

Os ramos dos organismos de vida livre (Rhabditophora + Catenulida).

Dentre os organismos de vida livre estão disponíveis 22 mitogenomas completos, sendo somente dois da classe Catenulida. Estes apresentam ordem gênica muito diferente aos demais representantes do filo Plathelminthes, além de inversões no sentido transcricional de alguns genes. Os dois mitogenomas de Catenulida da família Stenostomidea são semelhantes, com exceção da inversão completa do bloco *Nadh1*, *Cox3*, *Nadh6* e *CytB* entre eles (Figura 2.2).

Macrostomum lignano como único representante dos Macrostromorpha, acaba se posicionando como grupo irmão dos Catenulida (Figura 3.2), e com uma ordem gênica completamente distinta aos demais Platyhelminthes. Quando complementamos a análise Bayesiana com o organismo *Microstomum lineare*, também pertencente a Macrostromorpha, este fica incluído no grupo Rhabditophora.

A análise filogenética Bayesiana é concordante com os dados relativos ao arranjo da ordem gênica, desmembrando alguns clados, e sugere polifilia entre organismos dos grupos Stylochoidea (amarelo) e Leptoplanoidea (azul) (Figura 3.2). A ordem gênica do *Cotylea Stylochoplana maculata* é distinta a todos demais Plathelminthes (marcado com estrela). *Bothromesostoma personatum* é o único representante dos Rabdocoelas, assim isolado dos Polycladidas (parte inferior da figura) Tricladida (clado roxo), e com ordem gênica distinta a todos demais Plathelminthes.

Os Tricladida apresentam uma ordem gênica bem conservada, agrupando as duas superfamílias (Geoplanoidea e Planarioidea). A informação contida na história dos genes na análise Bayesiana corrobora com a ordem gênica (Figura 3.2).

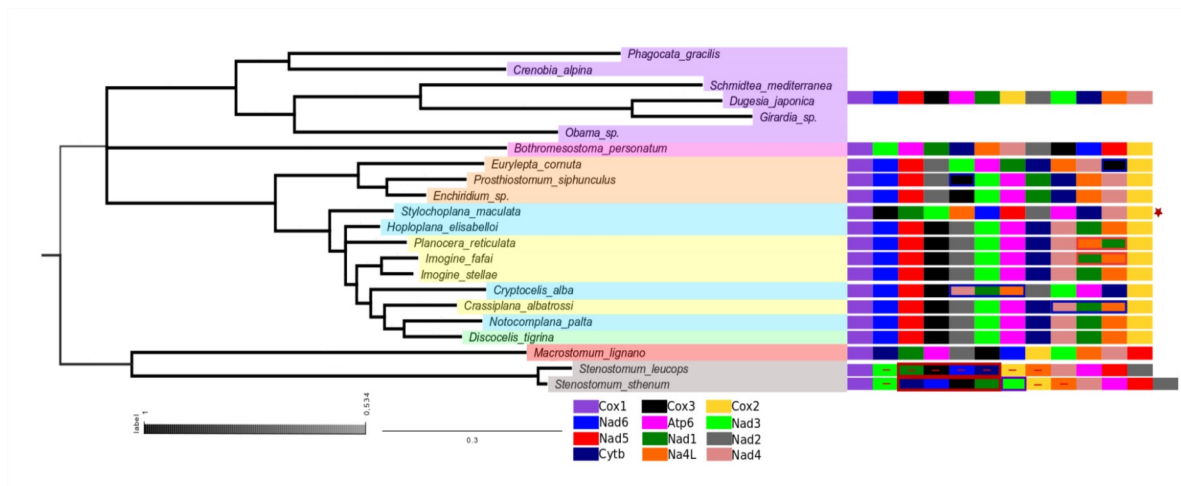


Figura 3.2: Corte da filogenia Bayesiana total das 12 CDSs, com somente os organismos de vida livre. Junto a representação filogenética com informação da ordem dos arranjos gênicos no mitogenoma representados por 12 quadros de cores diversificadas. Cada cor representa um gene. Os clados filogenéticos descritos estão representados por cores circundando os nomes das espécies. Catenulida cinza, Macrostomida vermelho, Ilyplonoidea verde, Leptoplanoidea azul, Stylochoidea amarelo, Cotylea bege, rosa Rabdocoela e roxo ordem Tricladida.

Análise filogenética dos Monogenia

Monogenia apresenta 31 espécies com mitogenomas completos depositados no NCBI. A informação filogenética obtida da sequência de aminoácidos dos 12 genes corrobora com duas grandes subclasses já descritas: Monopisthocotylea e Polyopisthocotylea (Figura 4.2). A subclasse Polyopisthocotylea é a primeira que ramifica dos Monogenia, posicionado como grupo irmão de todos os Monogenia Monopisthocotylea e demais Neodermatas (Figura 1.2). A ordem gênica mitocondrial dos Polyopisthocotylea também diverge em relação aos demais Neodermata e Monogenia Monopisthocotylea, pois há uma translocação do gene *Cox3*.

Os Monopisthocotylea com exceção de *Aglaiogyrodactylus forficulatus*, possuem uma ordem gênica uniforme para quase todos os demais clados parasitas. *Aglaiogyrodactylus forficulatus* fica ramificado mais distantes em relação aos demais da família *Gyrodactylidae*, corroborado pela ordem gênica, que apresenta uma translocação dos blocos *Nadh1* e *Nadh3*.

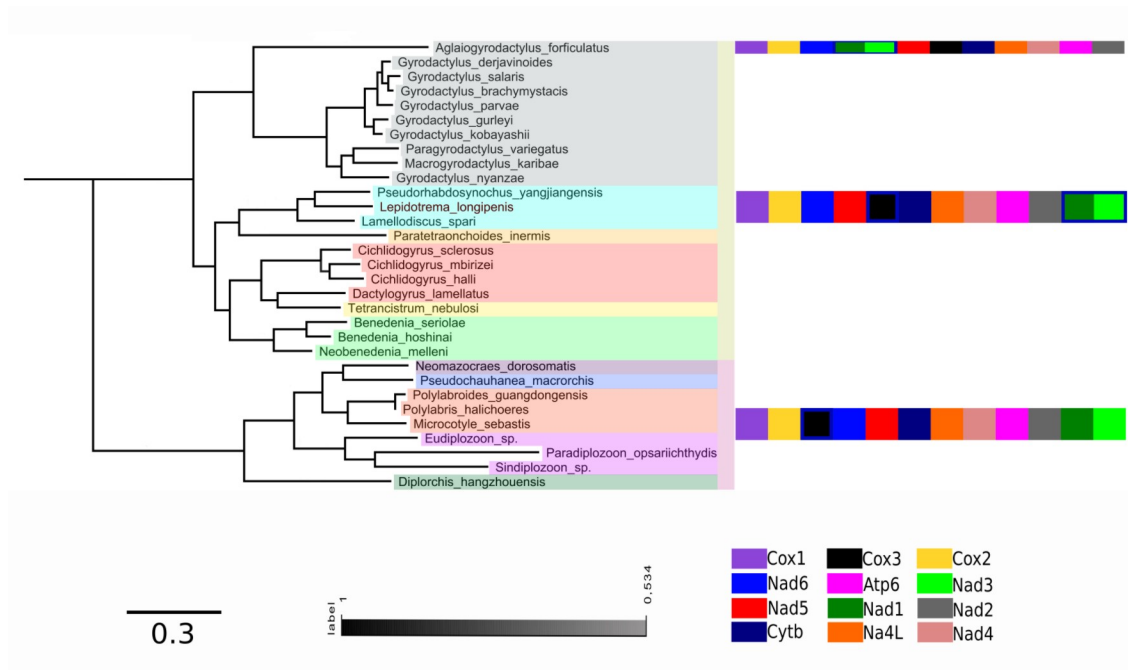


Figura 4.2: Filogenia da análise Bayesiana com os representantes Monogenea. Monopisthocotylea (barra superior verde) e Polyopisthocotylea (bege). Junto a representação filogenética com informação dos genes as ordens dos arranjos gênicos no mitogenoma representados por quadros de 12 cores diversificadas, cada cor representa um gene. Monopisthocotylea (barra superior/maior clado Monogenea): cinza: Gyrodactylidae, azul-claro: Diplectanidae, rosa: Dactylogyridea, amarelo: Ancyrocephalidae, pêssego: Tetraonchoididae e verde: Capsalidae. Polyopisthocotylea (clado inferior/menor): Roxo escuro: Mazocraeidae, Azul-marinho: Chauhaneidae, Microcotylidae: pêssego, Diplozoidae: roxo claro, Polystomatidae: verde-musgo.

Filogenética de Cestoda

Cestoda está representada por 53 mitogenomas depositados no NCBI, e é recuperada na análise filogenética Bayesiana, como grupo irmão de Trematoda. A família com mais representantes é *Taeniidae* com 23 mitogenomas, apresentando alta conservação na sequência de aminoácidos. A ordem gênica é a mesma para todos representantes do clado, e similar a quase todos Monogenea Monopisthocotylea e Trematodea, com exceção de alguns Schistosomatidae. Dentro da ordem dos *Caryophyllidea*, as famílias *Caryophyllaeidae* e *Lytocestidae* são parafiléticas entre si, mas sem nenhuma divergência na ordem gênica observado entre elas (Figura 5.2).

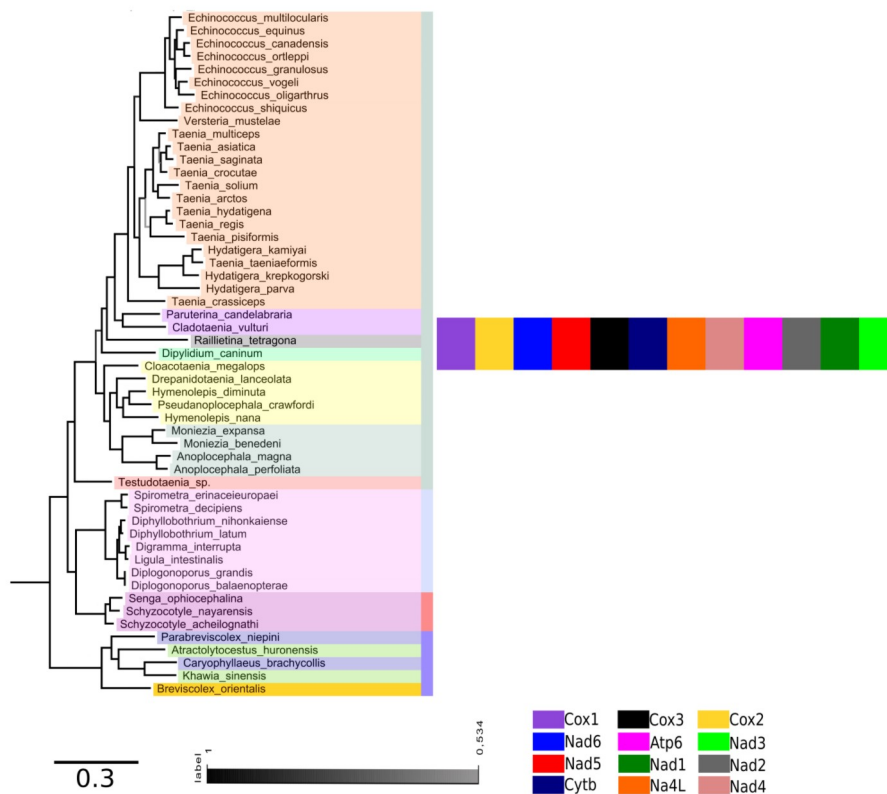


Figura 5.2: Ramificação da filogenia Bayesiana com os representantes Cestoda. Barras verticais são as Ordens: Cyclophyllidea (barra superior acinzentada), Diphyllbothriidea (azul-claro), Bothriocephalidea (pêssego) e Caryophyllidea (azul-marinho). Barras horizontais as famílias: Taeniidae (bege), Paruterinidae (roxo), Davaineidae (cinza), Dipylidiidae (verde), Hymenolepididae (amarelo), Anoplocephalidae (cinza claro), Proteocephalidae (pêssego), Diphyllbothriidae (rosa), Bothriocephalidae (roxo/rosa), Caryophyllaeidae (azul), Lytocestidae (verde) e Capingentidae (amarelo queimado). Junto a representação filogenética com informação dos genes as ordens dos arranjos gênicos no mitogenoma representados por 12 quadros de cores diversificadas, cada cor representa um gene.

Trematoda

A classe Trematoda é, entre os Plathelminthes, a que possui maior número de mitogenomas representados no NCBI, 59 deles completos. A família *Schistosomatidae* apresenta uma distinção em relação aos mitogenomas dos demais Neodermata, tanto em seu tamanho, que em vez da redução quando comparados aos organismos de vida livre, os seus mitogenomas costumam ter acima de 20Kb, com duplicações em CDSs, tRNAs e região rica em AT. A informação dos aminoácidos contido nos 12 genes coloca *Schistosomatidae* como grupo irmão dos demais Neodermata. A família *Schistosomatidae* também apresenta rearranjo gênico distinto aos demais Neodermata, com translocação do bloco *Atp6/Nadh2*, e uma inversão na ordem dos genes *Nadh3* e *Nadh1* (Figura 6.2).

A família *Diplostomatidae* forma parafilía com único representante da subfamília *Crassiphialinae* (*Posthodiplostomum centrarchi*) e os demais representantes da subfamília *Diplostominae*.

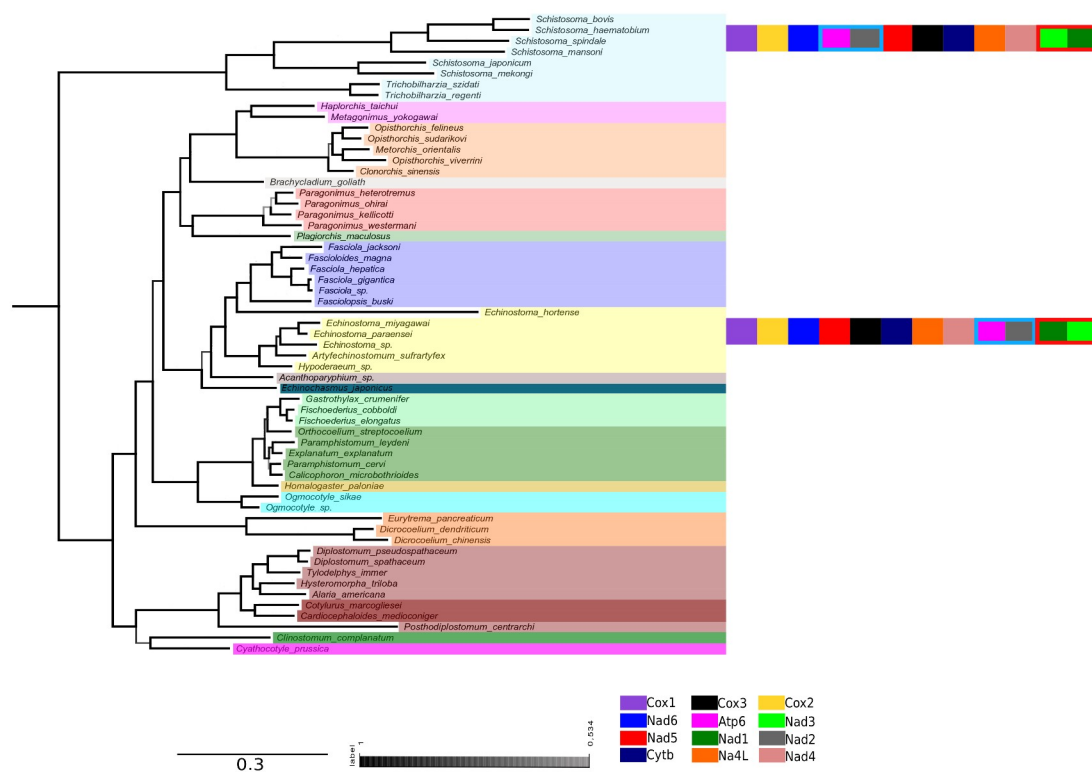


Figura 6.2: Ramificação da filogenia Bayesiana com os representantes Trematoda. Barras horizontais as famílias. Cyathocotylidae (Rosa), Clinostomidae (verde-limão), Diplostomatidae (marrom-claro), Strigeidae (marrom-escuro), Plagiorchioidea (pêssego), Notocotylidae (azul-ciano), Gastrodiscidae (amarelo-queimado), Paramphistomidae (verde-musgo), Gastrothylacidae (verde-claro), Echinochasmidae (verde-ciano), Himasthlidae (cinza), Echinostomatidae (amarelo-claro), Fasciolidae (roxo), Plagiorchiidae (verde-orvalho), Troglotrematidae (rosê), Brachycladiidae (cinza-claro), Opisthorchiidae (bege), Heterophyidae (rosa-claro), Schistosomatidae (azul-claro). Junto a representação filogenética com informação dos genes as ordens dos arranjos gênicos no mitogenoma representados por 12 quadros de cores diversificadas, cada cor representa um gene.

RNAs transportadores (tRNA)

Os mitogenomas de Platyhelminthes apresentam 22 tRNAs, e alguns destes podem estar presentes como múltiplas cópias. Os tRNAs em quase totalidade de cópias apresentam a mesma sequência e conformação estrutural. As sequências e estruturação são semelhantes em nível de família, e gênero em organismos de vida livre, mas altamente divergente entre as diferentes famílias. Entre os parasitas Neodermatas a sequência e suas estruturações são conservadas, mesmo em famílias e gêneros distantes. A codificação dos tRNAs de Catenulida diverge dos demais

Platyhelminthes em questão do tRNA Lisina (tRNA-K) (Figura 7.2). O tRNA-K em Catenulida é plesiomórfico aos demais invertebrados não Platyhelminthes sendo AAA (tRNA-UUU). A codificação do tRNA-K nos Platyhelminthes Rhabditophora e Neodermatas se dá por AAG (tRNA-CUU). A codificação AAA (tRNA-UUU) é Asparagina (Asn) em Platyhelminthes não Catenulida.

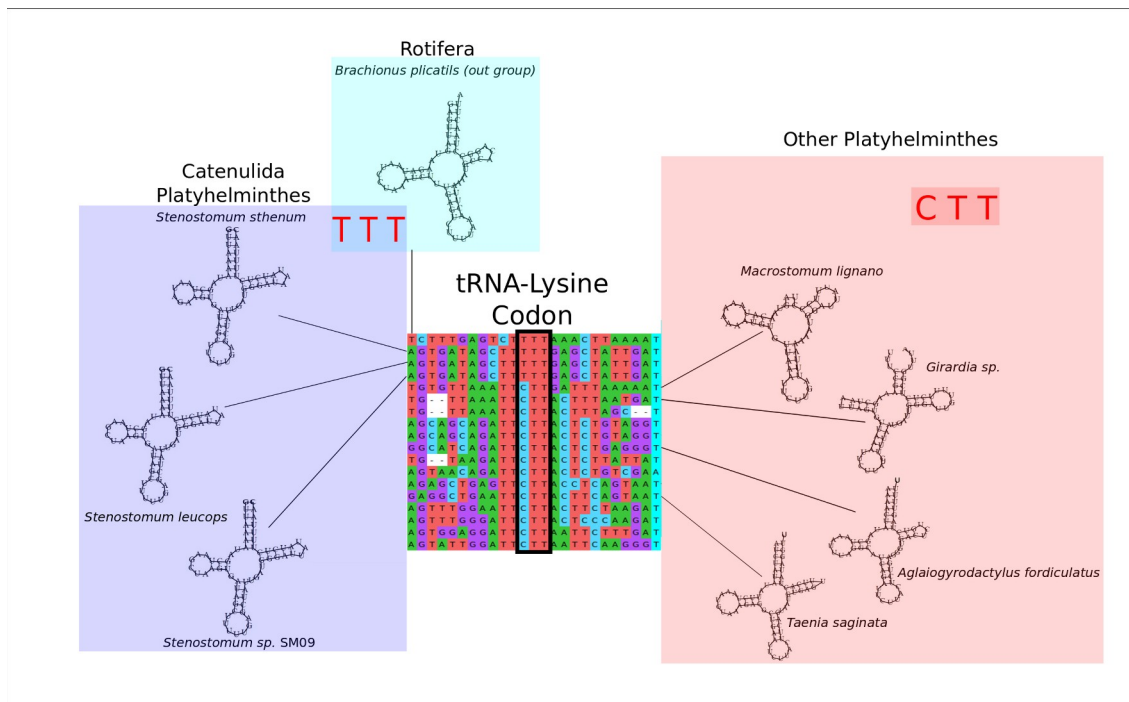


Figura 7.2: Conformação estrutural dos tRNA lysina (tRNA-K) em Platyhelminthes. Os blocos coloridos separam as diferentes formas estruturais do tRNA-K: roxo tRNA-K de Catenulida; em azul tRNA-K grupo externo, padrão nos demais invertebrados não Platyhelminthes; Vermelho os invertebrados Platyhelminthes, com representantes Rhabditophora e Neodermatas.

Discussão

A classe Catenulida ainda é uma grande questão relativo a sua posição filogenética no filo Platyhelminthes, em que se pergunta se ele é o filo basal, se grupo irmão dos Rhabditophora ou até mesmo se pertence aos Platyhelminthes. Estudos empregando marcadores moleculares como 18S, 28S, Cox1 recuperam a monofilia de Platyhelminthes e trazem Catenulida e Rhabditophora como grupos irmãos (Larsson e Jondelius 2008, Egger et al. 2015, Laumer et al. 2015). Os trabalhos de biologia do desenvolvimento, apontam para Rhabditophora e Catenulida também como grupos irmãos (Durán e Egger 2012). A filogenia mitocondrial trazida neste trabalho com 165 espécies de Platyhelminthes, recuperou a monofilia do filo, além de trazer Catenulida como grupo irmão dos demais Rhabditophora. O mesmo resultado é obtido com a matriz com 166 Platyhelminthes e a

inserção dos seis genes de *Microstomum lineare*. O grupo irmão se forma em ambas metodologias filogenéticas (Bayesiana e Verossimilhança).

As características mitocondriais do Catenulida *Stenostomum*, apresentam mais plesiomorfismo com invertebrados não Platyhelminthes que os seus parentes vermiformes. A codificação do aminoácido lisina no mitogenoma dos Catenulida também é plesiomórfico aos invertebrados externos aos demais Platyhelminthes (Telford et al. 2000). O plesiomorfismo se dá no códon AAA Lisina (Lys) em Catenulida e o uso do códon AAA Asparagina (Asn) em Rhabditophora (Telford et al. 2000). Os primeiros mitogenomas de Catenulida depositados no NCBI (*S. leucops* e *S. sthenum*), apontam para genes com codificação em ambos sentidos da sequência mitocondrial (Rosa et al. 2017 e Egger et al. 2017), outro carácter plesiomórfico a invertebrados não Platyhelminthes. Os Platyhelminthes Rhabditophora apresentam toda codificação genômica em único sentido. A ordem que os genes presentes no genoma mitocondrial em Catenulida se diferencia muito aos demais Platyhelminthes, não apresenta nenhum bloco comparativo com os Rhabditophora.

A classe Rhabditophora composta pelos Platyhelminthes de vida livre, apresenta genomas mitocondrias de apenas quatro ordens depositados no NCBI: *Macrostomida*, *Polycladida*, *Tricladida* e *Rhabdozoa*. O único representante com genoma mitocondrial completo/semi-completo de Macrostromorpha é *Macrostomum lignano* com 21 tRNAs em vez dos 22 (Egger et al. 2017), também não apresenta nenhum bloco conservado comparado aos demais Rhabditophora como em Catenulida. *Macrostomum lignano* quando comparado filogeneticamente, com bloco conservado das codificações dos 12 genes mitocondriais, acaba por ser mais próximo de Catenulida que com os demais Rhabditophora. A análise filogenética com *Microstomum lineare* e *Macrostomum lignano* com seis genes mitocondriais, coloca Macrostromorpha mais próximos dos demais Rhabditophora, em vez de Catenulida.

Os Polycladidas estão divididos em Cotylea e Acotylea e recuperaram suas monofilias (Kenny et al. 2018). Dentre os Acotylea apresentaram parafilias das superfamílias *Leptoplanoidea* e *Stylochoidea* (Kenny et al. 2018). A inserção de mais um organismo de outro trabalho da superfamília Stylochoidea, *Planocera reticulata* (NC_036051), também apresentou ser parafilético com *Crassiplana albatrossi*. A superfamília *Leptoplanoidea* apresentou parafilias entre seus representantes: *Stylochoplana maculata*, *Holoplana elisabelloi*, *Cryptocelis alba* e *Notocomplana palta*. O mitogenoma de *Stylochoplana maculata* apresenta uma ordem gênica com alta variação em relação aos demais Polycladidas. As variações e formações de parafilias, podem ser demonstradas tanto nas sequências de aminoácidos, como a ordem gênica de seus representantes.

O grupo dos Neodermatas recuperou sua monofilia, os grupos internos se formaram pela informação da sequência contida nos aminoácidos das CDSs. Os agrupamentos internos (ramos) também respeitaram as variações na ordem das CDSs. Os Neodermata formaram um padrão parafilético, o ramo da subclasse Polypisthocotylea apresentou ser de uma ramificação de um ancestral basal aos demais Neodermata (Monopisthocotylea, Monogenea, Cestoda e Trematoda).

A classe Monogenea está dividido em duas subclasses: Monopisthocotylea e Polyopisthocotylea. A filogenética molecular na maioria das análises, rejeita a monofilia de Monogenea, e coloca Polyopisthocotylea como basal aos demais Neodermatas (Justine 1998, Littlewood et al. 1999, Laumer and Giribet 2014). Polyopisthocotylea foi recuperado como grupo irmão de Monopisthocotylea e os demais Neodermata. Polyopisthocotylea apresenta uma modificação na posição da CDS *Cox3*, que diferencia dos Monopisthocotylea e demais Neodermata com uma ordem com blocos bem conservados. Monopisthocotylea apresenta a ordem gênica idêntica aos Cestoda e Trematoda, com exceção de *Aglaiogyrodactylus forficulatus*. O mitogenoma de *Aglaiogyrodactylus forficulatus* apresenta translocação do bloco *Nadh1* e *Nadh3*, gerando parafilia dentro da família *Gyrodactylidae*. A parafilia da família *Dactylogyridea* descrita por Mollaret et al. (2000) e Zhang et al. (2018), também foi observada em nosso estudo.

A classe Cestoda é monofilética, tendo como característica serem os primeiros vermes com segmentação externa e a proglotização interna. O genoma mitocondrial de Cestoda apresenta ordem gênica extremamente conservada, com arranjos somente em tRNAs (tRNA Leu (CUN) -tRNA Ser (UCN) -tRNA Leu (UUR) -tRNA Tyr) (Li et al. 2017). A ordem Caryophyllidea são considerados os primeiros eucestódios divergentes, e parasitas internos de peixes (Matoušková et al. 2019). A parafilia em Cestoda das famílias *Caryophyllaeidae* e *Lytocestidae*, já vinha sendo apontada na descrição do mitogenoma de *Parabreviscolex niepini* (Xi et al. 2018), esta foi novamente trazida neste estudo.

Os Trematoda apresentam parafilia na família *Diplostomatidae*, mas sem estruturação diferencial de ordem gênica, com um único representante da subfamília Crassiphialinae, esta topologia já foi apontada na descrição do mitogenoma de *Posthodiplostomum centrarchi* (Locke et al. 2018). A ramificação da superfamília *Schistosomatoidea* se distancia dos demais Trematoda, com a presença da translocação do bloco CDSs *Atp6* e *Nad2*, além da inversão na ordem *Nad3* e *Nad1*. Estas alterações já tinham sido apontadas em sequenciamento de *long pcr* de *Schistosoma mansoni* do Brasil comparado com espécies de *Schistosoma* da Ásia (Thanh et al. 2000).

Os mitogenomas dos Platyhelminthes são geralmente descritos com 15kb-20kb, compostos por 12 CDSs, 2 genes ribossomais e 22 genes de tRNAs. A geração de sequências *long reads* de alguns

organismos, por sequenciamento de DNA total ou por técnicas de *long pcr* (Le et al. 2020; Oey et al. 2019), tem apontado para genomas mais extensos, já demonstrado em mitogenomas de alguns *Schistosoma*. Os representantes de *Schistosoma* em que foi verificado que os mitogenomas expandidos usaram como metodologia de *long PCR* ou sequenciamento Nanopore (Kinkar et al. 2021). Já os trabalhos que não usaram destas técnicas e sim o sequenciamento de *short reads* (Illumina) recuperaram mitogenomas de 14 Kb-15kb para estas espécies. Os mitogenomas de *S. mediterranea* apresentam variações relacionadas a sua sexualidade, na forma sexual com 27,133 pb e assexual 17,176 pb (Ross et al. 2016). O sequenciamento Nanopore de *S. leucops* apresenta como sequência consenso um mitogenoma de + 27Kb, com repetição de um bloco inteiro contendo genes de tRNAs, ribossomais e CDSs.

O sequenciamento *long reads* pode ser uma importante ferramenta para entendermos a evolução mitocondrial dos Platyhelminthes, e suas estruturações mitocondriais. Uma ferramenta para evitar montagens estruturais errôneas, geradas por haver mais duplicações ao longo do mitogenoma. O mitogenoma de *S. leucops* depositado no NCBI apresenta sequência de aminoácidos correta, mas sua estruturação não condizente com *S. sthenum* e nem com seu sequenciamento *long read*. A estruturação pode ter sido equivocada, devido a utilização de *reads* curtos Illumina e montador usado. Algumas famílias de Platyhelminthes estão pouco representadas nos bancos de dados, principalmente os organismos de vida livre, e os dados existentes levam a parafilias de algumas famílias. Os Rhabditophora apresentam uma instabilidade na ordem gênica muito maior que nos Neodermata. Organismos que compartilham de mesma família dentro de Rhabditophora, possuem ordem gênica distinta, por outro lado, representantes de famílias diferentes, ordens gênicas semelhantes. Algumas famílias distintas com ordens semelhantes, ficaram filogeneticamente próximas. As codificações e translocação de ordem em regiões CDSs já é um sinal para separação de ramos nos Neodermata em que a ordem gênica também apresenta ser quase que em sua totalidade conservada.

Referência:

Hyman, L.H. (1951). *The Invertebrates: platyhelminthes and Rhynchocoela*.

The Acoelomate Bilateria. McGraw- Hill, New York, V. 2 pp 550.

Park, J.K., Kim, K.H., Kang, S. et al.(2007). A common origin of complex life cycles in parasitic flatworms: evidence from the complete mitochondrial genome of *Microcotyle sebastis* (Monogenea: Platyhelminthes). *BMC Evol. Biol.* V. 7 pp 11. <https://doi.org/10.1186/1471-2148-7-11>

Jondelius, U., Ruiz-Trillo, I., Bagnà, J. and Riutort, M. (2002), The Nemertodermatida are basal bilaterians and not members of the Platyhelminthes. *Zoologica Scripta*, V. 31 pp 201-215. <https://doi.org/10.1046/j.1463-6409.2002.00090.x>

Telford, M.J., Lockyer, A.E., Cartwright-Finch, C., Littlewood, D.T.J. (2003). Combined large and small subunit ribosomal RNA phylogenies support a basal position of the acoelomorph flatworms. *Proc. R. Soc. London.* V. 270 pp. 77-1083.

Wallberg, A., Curini-Galletti, M., Ahmadzadeh, A. and Jondelius, U. (2007). Dismissal of Acoelomorpha: Acoela and Nemertodermatida are separate early bilaterian clades. *Zool. Scripta*. V. 36 509–523.

Larsson, K., Ahmadzadeh, A., Jondelius, U. (2008). DNA taxonomy of Swedish Catenulida (Platyhelminthes) and a phylogenetic framework for catenulid classification. *Org. Divers. Evol.*, V.8 pp. 399-412.

Egger, B, Lapraz, F., Tomiczek, B., et al. (2015). A transcriptomic-phylogenomic analysis of the evolutionary relationships of flatworms. *Curr Biol.* V. 25 pp1347-1353. [doi:10.1016/j.cub.2015.03.034](https://doi.org/10.1016/j.cub.2015.03.034)

Laumer, C.E., Andreas, E., Gonzalo, G., Sánchez A.A., (2015). Nuclear genomic signals of the 'microturbellarian' roots of platyhelminth evolutionary innovation. *Elife*. V.4 pp 1-31. <https://doi.org/10.7554/eLife.05503>.

Schoch, C.L., Ciufo, S., Domrachev, M., et al. (2020). NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database (Oxford)*. doi: 10.1093/database/baaa062.: 32761142; PMID: PMC7408187.

Noreña, C., Damborenea, C., Brusa, F. (2005). A taxonomic revision of South American species of the genus *Stenostomum* O. Schmidt (Platyhelminthes: Catenulida) based on morphological characters. *Zool. J. Linn. Soc.* V.144 pp 37–58.

Sopott-Ehlers, B., Ehlers, U. (1999). Ultrastructure of spermiogenesis and spermatozoa in *Bradynectes sterreri* and remarks on sperm cells in *Haplopharynx rostratus* and *Paromalostomum fuscum*: phylogenetic implications for the Macrostromorpha (Plathelminthes). *Zoomorphology* V.119 pp105–115. <https://doi.org/10.1007/s00435005008>

Larsson, K., & Jondelius, U. (2008). Phylogeny of Catenulida and support for Platyhelminthes. *Organisms Diversity & Evolution*, V. 8 pp 378-387.

Morrow, E.H. (2004), How the sperm lost its tail: the evolution of aflagellate sperm. *Biological Reviews*, 79: 795-814. <https://doi.org/10.1017/S1464793104006451>

Willems, M., Leroux, F., Claeys, M., Boone, M., Mouton, S., Artois, T. and Borgonie, G. (2009), Ontogeny of the complex sperm in the macrostomid flatworm *Macrostomum lignano* (Macrostromorpha, Rhabditophora). *J. Morphol.*, 270: 162-174. <https://doi.org/10.1002/jmor.10675>

Egger, B., Bachmann, L. & Fromm, B. Atp8 is in the ground pattern of flatworm mitochondrial genomes. *BMC Genomics* 18, 414 (2017). <https://doi.org/10.1186/s12864-017-3807-2>

Kearn, G.C.(1994). Evolutionary expansion of the Monogenea. *International Journal for Parasitology*. V.24 pp 1227-1271. [https://doi.org/10.1016/0020-7519\(94\)90193-7](https://doi.org/10.1016/0020-7519(94)90193-7).

Kmentová, N., Gelnar, M., Mendlová, M. et al. (2016). Reduced host-specificity in a parasite infecting non-littoral Lake Tanganyika cichlids evidenced by intraspecific morphological and genetic diversity. *Sci Rep*. V.6 pp 39605. <https://doi.org/10.1038/srep39605>

Telford, M.J., Herniou, E.A., Russell, R.B. and Littlewood, T.J. (2000). Changes in mitochondrial genetic codes as phylogenetic characters: two examples from the flatworms. *Proceedings of the National Academy of Sciences, United States of America* V. 97: pp 11359-11364.

Justine, J.L. (1998). Non-monophyly of the monogeneans? *International Journal for Parasitology*. V.28 pp 1653–1657. [https://doi.org/10.1016/S0020-7519\(98\)00060-5](https://doi.org/10.1016/S0020-7519(98)00060-5)

Littlewood, D.T., Rohde, K., Bray, R.A., Herniou, E.A. (1999). Phylogeny of the Platyhelminthes and the evolution of parasitism. *Biological Journal of the Linnean Society*. V.68 pp 257–287. <https://doi.org/10.1111/j.1095-8312.1999.tb01169.x>

Mollaret, I., Jamieson, B. G., Justine, J.L. (2000). Phylogeny of the Monopisthocotylea and Polyopisthocotylea (Platyhelminthes) inferred from 28S rDNA sequences. *International Journal for Parasitology*. V.30 pp 171-185.

Zhang, D., Li, W.X., Zou, H. et al. (2018). Mitochondrial genomes of two diplectanids (Platyhelminthes: Monogenea) expose parafphyly of the order Dactylogyridea and extensive tRNA gene rearrangements. *Parasites Vectors* V.11 pp 601. <https://doi.org/10.1186/s13071-018-3144-6>

Xi, B-W., Zhang, D., Li, W-X., Yang, B-J., Xie, J. (2018). Characterization of the complete mitochondrial genome of *Parabreviscolex niepini* Xi et al., 2018 (Cestoda, Caryophyllidea). *ZooKeys*. V.783 pp 97-112. <https://doi.org/10.3897/zookeys.783.24674>

Locke, S.A., Van Dam, A.R., Caffara, M., Pinto, H.A., López-Hernández, D. and Blanar, C.A. (2018). Validity of the Diplostomoidea and Diplostomida (Digenea, Platyhelminthes) upheld in phylogenomic analysis. *International Journal for Parasitology*. V.48 pp 1043–1059. [10.1016/j.ijpara.2018.07.001](https://doi.org/10.1016/j.ijpara.2018.07.001)

Gissi, C., Iannelli, F. & Pesole, G. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity* 101, 301–320 (2008). <https://doi.org/10.1038/hdy.2008.62>

Elson, J.L., Lightowers, R.N. (2006). Mitochondrial DNA clonality in the dock: can surveillance swing the case?. *Trends Genet*. V. 22 pp 603-7. doi: 10.1016/j.tig.2006.09.004. Epub 2006 Sep 18. PMID: 16979783.

Lavrov, D.V., Lang, B.F. (2005) Transfer RNA gene recruitment in mitochondrial DNA. *Trends in Genetics*. V. 21 pp 129–133.

Boore J.L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, V. 27 pp 1767–1780. <https://doi.org/10.1093/nar/27.8.1767>

Monnens, M., Thijs, S., Briscoe, A.G., et al., (2020). The first mitochondrial genomes of endosymbiotic rhabdocoels illustrate evolutionary relaxation of *atp8* and genome plasticity in flatworms. *International Journal of Biological Macromolecules*. <https://doi.org/10.1016/j.ijbiomac.2020.06.025>

Kinkar, L., Robin, B.G., Bonnie, L.W., David, R., Littlewood, D.T.J. , Bill, C.H. Chang, Andreas J.S., Pasi K.K., and Neil D. Y. (2021). "Nanopore Sequencing Resolves Elusive Long Tandem-Repeat Regions in Mitochondrial Genomes" *International Journal of Molecular Sciences*. V. 22, no. 4 pp 1811. <https://doi.org/10.3390/ijms22041811>

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol*. V.35 pp 1547-1549. doi: 10.1093/molbev/msy096. PMID: 29722887; PMCID: PMC5967553.

Talavera, G., Castresana, J., (2007). Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Systematic Biology*. V. 56 pp 564–577. <https://doi.org/10.1080/10635150701472164>

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., ... Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*. V. 29 pp 644–652. <https://doi.org/10.1038/nbt.1883>

Donath, A., Jühling, F., Al-Arab, M., Bernhart, S.H., Reinhardt, F., Stadler, P.F., Middendorf, M., Bernt, M. (2019). Improved annotation of protein-coding genes boundaries in metazoan mitochondrial genomes *Nucleic Acids Research*. V. 47 pp 10543–10552.

Le, T., Pham, L., Doan, H., Le, X., Saijuntha, W., Rajapakse, R., and Lawton, S. (2020). Comparative mitogenomics of the zoonotic parasite *Echinostoma revolutum* resolves taxonomic relationships within the ‘*E. revolutum*’ species group and the Echinostomata (Platyhelminthes: Digenea). *Parasitology*. V. 147 pp 566-576. doi:10.1017/S0031182020000128

Oey, H., Zakrzewski, M., Narain, K., Devi, K.R., Agatsuma, T., Nawaratna, S., Gobert, G.N., Jones, M.K., Ragan, M.A., McManus, D.P., Krause, L. (2019). Whole-genome sequence of the oriental lung fluke *Paragonimus westermani*. *GigaScience*. V. 8. <https://doi.org/10.1093/gigascience/giy146>

Ross, E., Blair, D., Guerrero-Hernández, C., & Sánchez Alvarado, A. (2016). Comparative and Transcriptome Analyses Uncover Key Aspects of Coding- and Long Noncoding RNAs in Flatworm Mitochondrial Genomes. *G3* (Bethesda, Md.). V. 6pp 1191–1200. <https://doi.org/10.1534/g3.116.028175>

Kenny, N.J., Noreña, C., Damborenea, C., Grande, C. (2019). Probing recalcitrant problems in polyclad evolution and systematics with novel mitochondrial genome resources. *Genomics*. V. 111 pp 343-355. <https://doi.org/10.1016/j.ygeno.2018.02.009>.

Li, W.X., Zhang, D., Boyce, K. et al. (2017). The complete mitochondrial DNA of three monozoic tapeworms in the Caryophyllidea: a mitogenomic perspective on the phylogeny of eucestodes. *Parasites Vectors*. V.10 pp 314. <https://doi.org/10.1186/s13071-017-2245-y>

Matoušková, M., Bílý, T., Bruňanská, M. et al. (2019). New data on spermiogenesis and trepaxonematan axoneme in basal tapeworms (Cestoda, Caryophyllidea, Lytocestidae) parasitizing cyprinid fishes. *Sci Rep*. V. 9 pp12881. <https://doi.org/10.1038/s41598-019-49312-9>

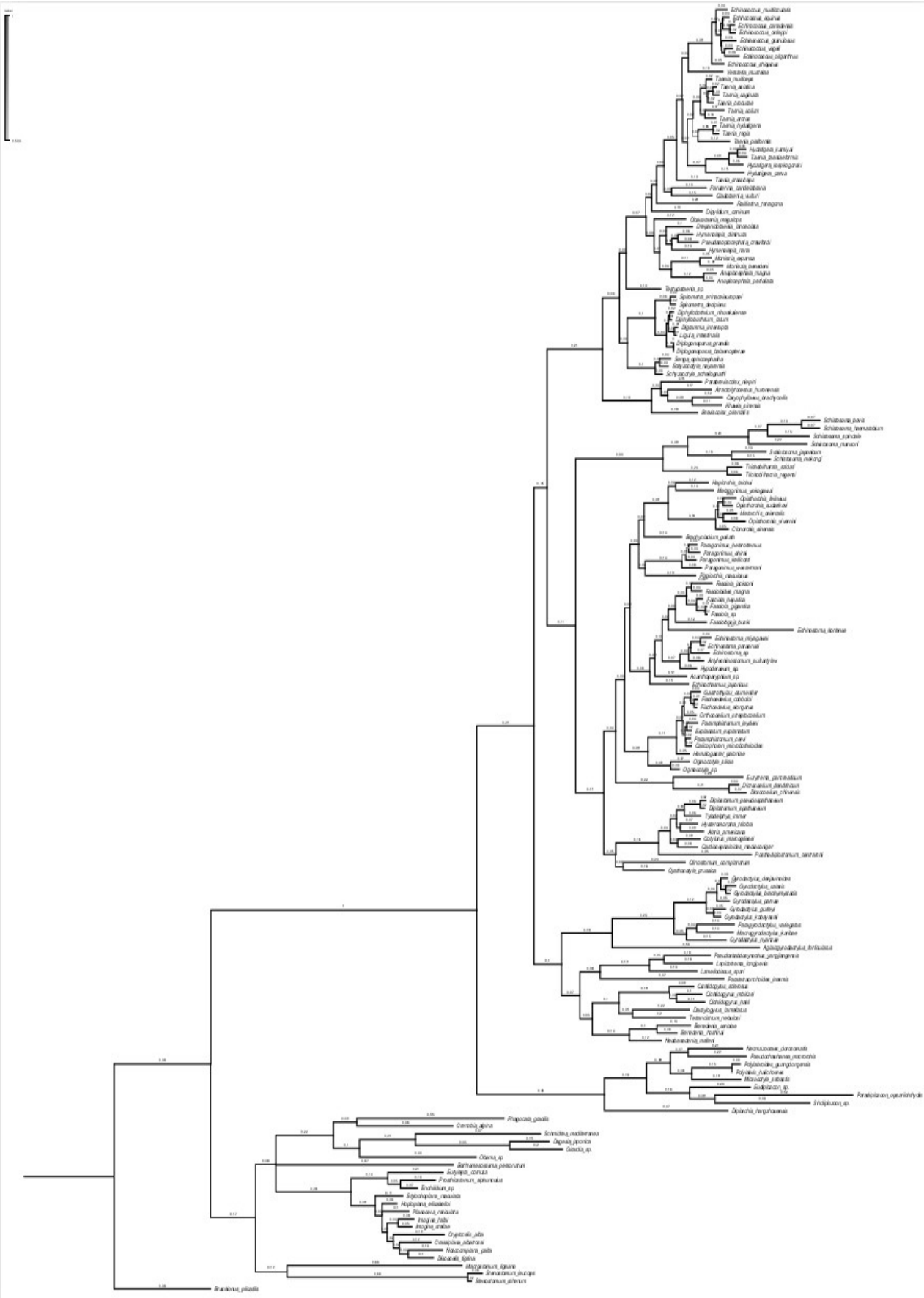
Thanh, H.L., Blair, D., Agatsuma, T., et al. (2000). Phylogenies Inferred from Mitochondrial Gene Orders—A Cautionary Tale from the Parasitic Flatworms, *Molecular Biology and Evolution*. V. 17 pp1123–1125. <https://doi.org/10.1093/oxfordjournals.molbev.a026393>

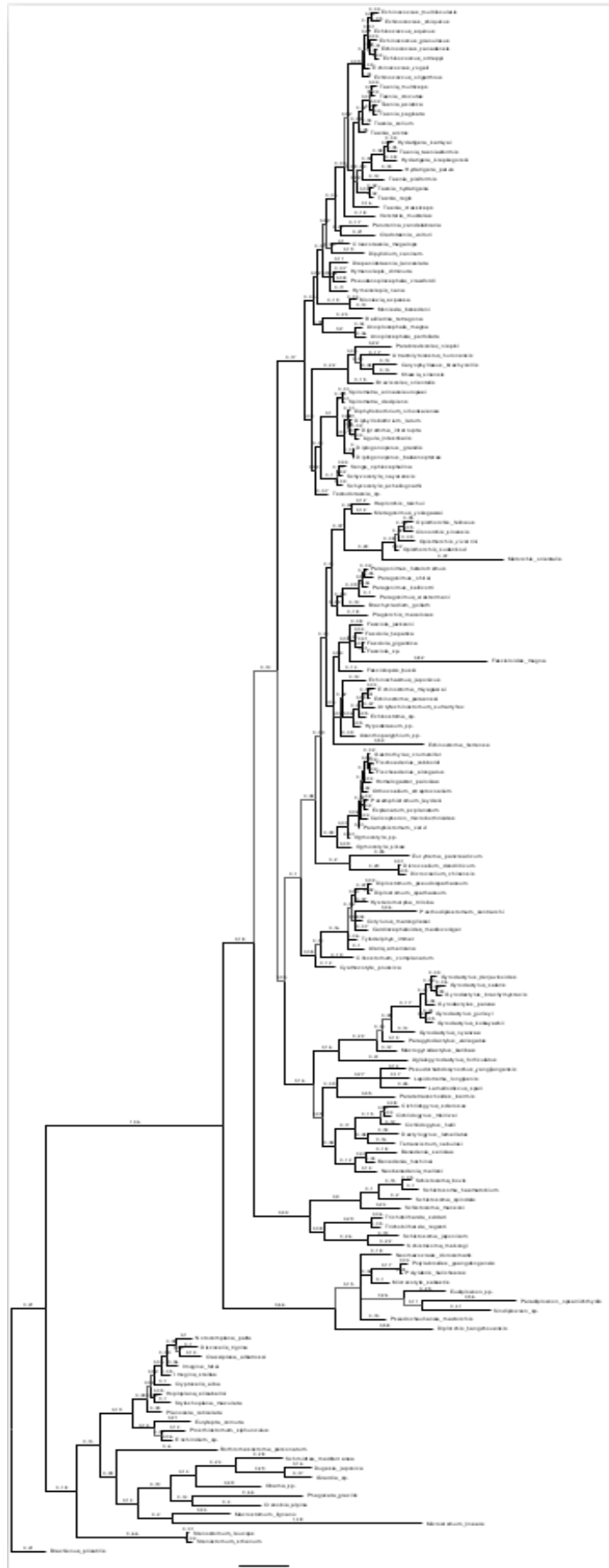
Material suplementar

*Material supl. 1.2: Árvore filogenética originada da matriz com 165 Platyhelminthes com suas 12 CDSs. (link: https://www.dropbox.com/s/khk94bnm5vopbhu/filo_165_12CDS.pdf?dl=0)

*Material supl. 2.2: Árvore filogenética originada da matriz com 166 Platyhelminthes com suas seis CDSs. (link: https://www.dropbox.com/s/rnb21h76vaepd9i/filo_166_6CDS.pdf?dl=0)

*Tabela suplementar 1 (anexo da tese “tabela_mitos_platy”) (link: https://www.dropbox.com/scl/fi/110pi08xe8wvhi77655dr/tabela_platy_mitos.ods?dl=0&rlkey=gaf7mm36a7eyxde2fnox1kflx)





Capítulo 3

Revisiting the regeneration of *Stenostomum leucops* (Catenulida, Platyhelminthes)



Revisiting the regeneration of *Stenostomum leucops* (Catenulida, Platyhelminthes)

Marcos T. Rosa & Elgion L. S. Loreto

To cite this article: Marcos T. Rosa & Elgion L. S. Loreto (2021): Revisiting the regeneration of *Stenostomum leucops* (Catenulida, Platyhelminthes), *Invertebrate Reproduction & Development*, DOI: [10.1080/07924259.2021.1994475](https://doi.org/10.1080/07924259.2021.1994475)

To link to this article: <https://doi.org/10.1080/07924259.2021.1994475>



Published online: 26 Oct 2021.



Submit your article to this journal [↗](#)



Article views: 15




View related articles [↗](#)



View Crossmark data [↗](#)



Revisiting the regeneration of *Stenostomum leucops* (Catenulida, Platyhelminthes)

Marcos T. Rosa^a and Elgion L. S. Loreto^b 

^aPrograma de Pós-Graduação em Biodiversidade Animal, Universidade Federal de Santa Maria, Santa Maria, Brazil; ^bDepartment of Biochemistry and Molecular Biology, Universidade Federal de Santa Maria, Santa Maria, Brazil

ABSTRACT

Stenostomum are small flatworms that live in freshwater and normally reproduce by paratomy. They are basal in Platyhelminthes phylogeny. For more than a century, species of this genus, especially *S. leucops*, have been used in regeneration studies. However, some contradictory results remain, as there are concerns about which body parts regenerate, and their routes of regeneration. Problems in species identification are a possible explanation for these contradictory results. Regeneration studies were repeated, transforming the worms with vectors carrying an eye-specific GFP (green fluorescence protein) marker in a characterized strain. All regeneration patterns described previously were observed in this strain, showing that these complex regeneration routes are probably associated with the different parts of the body that were cut. *Stenostomum* species are promising experimental organisms for using a transcriptome approach in regeneration studies.

ARTICLE HISTORY

Received 26 July 2021
Accepted 12 October 2021

KEYWORDS

Regeneration; stem cells; planarians; green fluorescence protein

Introduction

The class Catenulida (Platyhelminthes) is represented by approximately 100 described species, are small freshwater flatworms, and generally range in length from 0.5 to 5 mm (Noreña et al. 2005). They normally reproduce by paratomy, an asexual reproductive process in which development of structures typical of the anterior region in posterior parts of body occur (Rieger 1986; Noreña et al. 2005). New worms developing along the body axis is called a zooid, and in some species, chains with several zooids can occur (Hyman 1951). After complete development fission of each zooid perpendicular to the anteroposterior axis occurs, forming new worms. One of the most studied Catenulida is *Stenostomum leucops* (Dugès 1828). This species is cosmopolitan and easily maintained in culture, where they reproduce quickly, regenerate extensively, and have internal organs that can be readily observed (Nuttycombe and Waters 1938). These characteristics make it an excellent experimental organism.

For more than a century, *Stenostomum*, especially *S. leucops*, has been used in studies on regeneration. However, these studies have been marked by contradictory results. Ritter and Congdon (1900) described that if a worm, formed by two or more zooids, was sectioned in such a way that a portion containing the cephalic structures, in formation, was positioned in the posterior

part of the sectioned worm, these cephalic structures migrate to an anterior position of the body, to constitute the cephalic part of the regenerated worm. Child (1902) did not observe migration of these structures; instead, a reduction and resorption of the ganglionic masses in the posterior position of a regenerating worm was reported, depending on the developmental stage of these structures. Resorption is complete in a young cephalic structure. However, according to Child (1902), resorption does not occur when the structures are well developed. Other contradictory results concern whether the anterior and posterior halves of a zooid can be reconstituted completely. Hartmann (1922) and Ruhl (1927) described that both parts regenerate, while Ritter and Congdon (1900) and Child (1902) reported that only the anterior part is regenerated. Van Cleave (1929), in a comparative study about regeneration in *S. grande*, *S. tenuicauda*, and *S. leucops*, found that the latter species differs from the former two because regeneration was less common, slower, and resorption of the cephalic structures did not occur. However, Van Cleave (1929) and Nuttycombe and Waters (1938) indicated that at least part of these contradictory results could be associated with problems of species identification.

The validity of *S. leucops* as a species has been questioned by Nuttycombe and Waters (1938) and Marcus (1945). For the first authors '... it seems likely to us that

the descriptions of *S. leucops* are so unreliable as to invalidate completely that species in any fauna; we rather strongly suspect that is a synonym for *Stenostomum*'. Furthermore, it seems that in early studies, many species were used as *S. leucops*. However, further studies done by Borkott (1970) and Noreña et al. (2005) validated this species. More recently, Larsson et al. (2008) and Yamazaki et al. (2012) used DNA sequences of 18 S rDNA and COI to show that *S. leucops* consists of a monophyletic group. Nevertheless, phylogenetic studies using different genes suggest that this taxon corresponds to a species complex (Larsson et al. 2008; Rosa et al. 2015). A strain of *S. leucops*, having an average size of 0.98 ± 0.09 mm for worms with two zooids, was characterized. The time spent in paratomy is approximately 42.5 h at 28°C, producing zooids containing approximately 2400 cells. Typically, the worms have two zooids, but it is a plastic phenotypic trait, and up to five zooids can sometimes be observed depending on the environmental conditions (Rosa et al. 2015). Additionally, this strain can express the *gfp* gene when it is carried on plasmids in its microbiome (Rosa and Loreto 2019).

Planarians are a powerful model for studies of regeneration, stem cells, and ageing due to their regenerative properties, combined with the development of new molecular technologies and sequencing. *Schmidtea mediterranea*, *Dugesia (Girardia) japonica*, and *D. (G.) tigrina* are the species most commonly used in planarian research (Oviedo et al. 2008; Sánchez Alvarado 2012). Due to the basal position of Catenulida in Platyhelminthes phylogeny (Rohde et al. 1993; Campos et al. 1998), members of this order can be a powerful tool for comparative studies with planarians. *Stenostomum leucops*, which is very easily maintained, could be a candidate as an organism for this comparative study.

In this study, the regeneration pattern was characterized in sectioned worms using the *S. leucops* strain as previously described. Green fluorescent protein (GFP) expression in the ciliated pit was used as a marker for cephalic regeneration, degeneration, or migration. Different regeneration routes were observed, similar to those described previously.

Materials and methods

Animals

Worms were collected in a dam at the Federal University of Santa Maria, Santa Maria, Brazil (53°17' W; 29°28' S) in March 2009. The culture was initiated from a single worm (clonal isolate) and was called SL01-sm01. Since then, a stock has been maintained at 28°C, in a 25 ml

culture flask with 10 ml of reconstituted water (Knakievicz et al. 2006). Every 3 days, half of the water volume was changed and about 2 mg of powdered milk was added to the medium. The specimens were identified following Noreña et al. (2005) and by DNA barcoding (Rosa et al. 2015). A detailed description of this strain and its biological characteristics can be found in Rosa et al. (2015).

Regeneration assay

The worms were cut with glass needles, made of capillary tubes, by hand-pulling these capillaries over the flame of an alcohol lamp. The worms were selected in the two-zooid phase, placed on a slide with a drop of 0.5 µl reconstituted water, then viewed under an inverted microscope at 100× magnification. The worms were sectioned following two different section schemes, although in both, the worms were sectioned in three parts (Figure 1). In the first scheme, one cut was in the middle of the pharyngeal gland region, and the other cut was positioned after the developing fission plane and the ciliated pits that develop on half of the body (red lines in Figure 1). In this case, the anterior animal portion does not contain the regions with the intestine and Minot cells. However, the median and posterior sections contain these regions. In the second scheme, the cuts were performed in the cephalic and median regions containing the intestine and Minot cells, but in the posterior portion these cells were absent (blue lines in the Figure 1). The rationale for this sectioning scheme was to test the hypothesis that most stem cells are in the regions containing Minot cells. For each sectioned animal the three parts were maintained in a drop of water on the slide in a wet chamber, at 28°C, and were observed with an inverted microscope every 3 h until complete regeneration or death. The sample size of the regeneration assays was 100 worms. Differences in the regeneration classes were tested by χ^2 .

GFP expression

Previously, we demonstrated that *S. leucops* has stable expression of GFP, mainly in the ciliated pits, when the *Gfp* gene is carried in its microbiome (Rosa and Loreto 2019). SL01-sm01, which was the strain used, holds *Escherichia coli* in its microbiome, and this property was used to induce GFP expression, transforming the microbiome by electroporation. The GFP mark in the ciliated pit was used in this study to help identify the cephalic region.

For the genetic transformation assays, plasmids pBac [3xp3-EGFPafm] (Horn et al. 2000), containing the inverted terminal repeats (ITRs) of the *piggyBac*

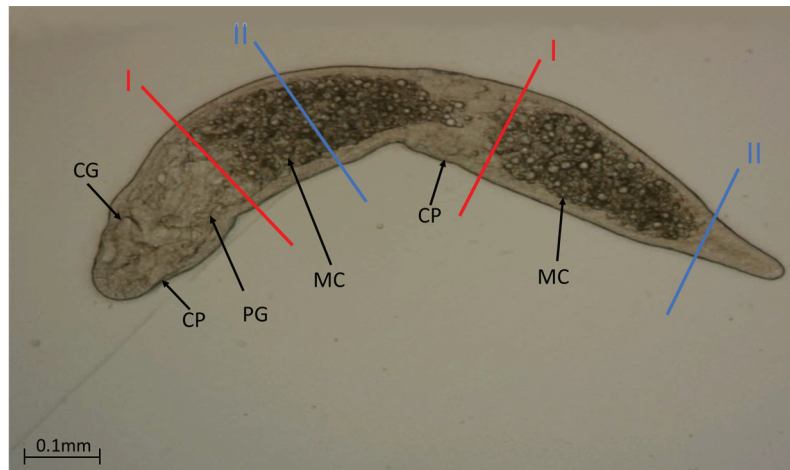


Figure 1. *Stenostomum leucops* with zooid formation in the regeneration assay. Red lines represent sectioning lines in the first section schemes and the blue lines the second scheme. In both schemes the body is sectioned in three parts: 1) anterior cephalic part; 2) median portion and 3) posterior body part. CP = ciliated pits; CG = cerebral ganglia; PG = pharyngeal glands; MC = Minot cells.

transposon, and the GFP gene marker, under the control of an artificial promoter containing three Pax-6 homodimer binding sites, were used. This promoter drives strong expression of the GFP protein in the eye tissues (Gehring 2002). The worms were electroporated by a single pulse of 80 V for 1 msec using a Gene Pulser Xcell Electroporation System (BIO RAD)

apparatus. Fifty worms were placed in a 2 mm cuvette with 200 μ l reconstituted water, 35 μ l PBS and 1.5 μ g of each plasmid. The experiment was replicated four times. After electroporation, the worms were rinsed in reconstituted water and maintained individually in wells of plastic cell culture plates with reconstituted water (3 ml per well) and fed

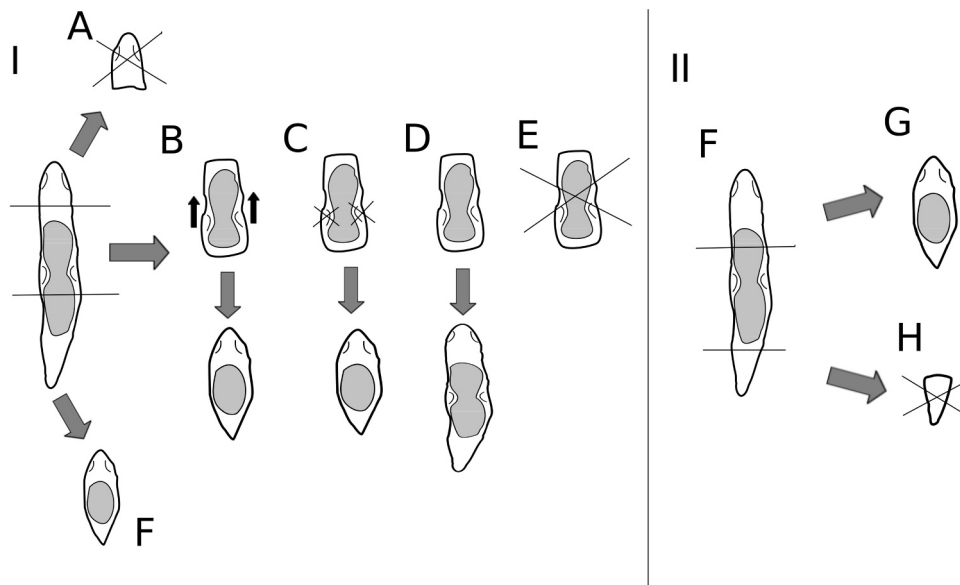


Figure 2. For the regeneration assay, worms were cut into three parts. In the first section scheme (i), the anterior part (a) never regenerates. The posterior part always regenerates (F). For the median part, four different routes can be taken: (b) the head structures that are developing in the fission plane can move to the anterior part and cause regeneration; (c) those structures can degenerate, and new ones arise in the anterior region; (d) new structures arise in the anterior part, and a posterior part is also regenerated, forming a worm with two zooids; (e) after some days without differentiation, the structures degenerate. In the second section scheme (II), when in the anterior part contains part of the intestine (f), its part regenerates (g); if the posterior part does not receive part of the intestine, it is not able to regenerate (h).

with milk powder. As a control, worms were exposed to the same solutions and plasmids but did not receive the electric pulse.

Results

After mechanical sectioning, diverse routes of regeneration were observed in the worms. The specific route appears to depend on the planes of fission and the developmental stage of the zooids (Figure 2). No regeneration was observed when the product of section had parts of the animal that excluded regions containing intestinal portion or Mitot cells (Figure 2 I-A; II-H). However, if some intestinal portion was present, regeneration was normally

observed (Figure 2: I-B, C, D, and F; II-G). Several regeneration patterns were observed for the median body portion when the zooid's cephalic structures were maintained in the posterior portion of the fissioned body. For the first section scheme, a possible outcome was the movement of the zooid's ciliated pits and cephalic structures in development to the anterior region (Figures 2(i-b) and Figure 3(a)). Another observed outcome was the degeneration of the posterior ciliated pits and cephalic structures with the formation of new ones in the anterior region (Figures 2(i-c) and Figures 3(b)). In some cases, a new 'head' was formed in the anterior region, forming two zooids that proceeded in paratomy (Figures 2(i-d) and Figures 3(c)). Finally, some worms

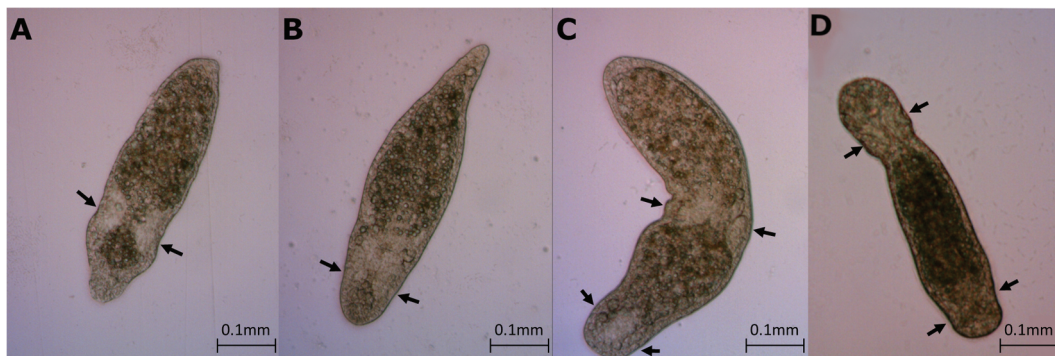


Figure 3. Regeneration patterns in the median portion of *Stenostomum leucops* at 24 h after ablation observed in the first and second section schemes. . The arrows point to the ciliated pits. (a) migration from the posterior cephalic region to the anterior region of the body; (b) degeneration of the posterior cephalic structures and formation of a new anterior cephalic region; (c) formation of a new anterior cephalic region and body region behind the posterior cephalic region, followed by paratomy; (d) Worms that died without showing any regenerative pattern in less than 24 h.

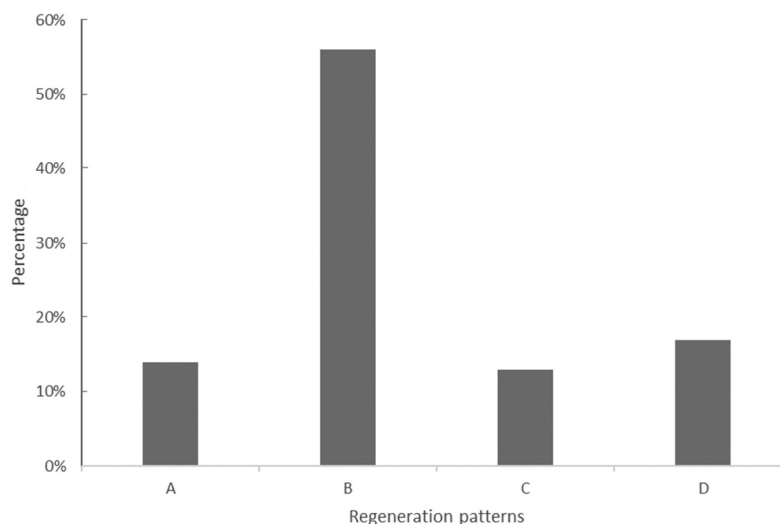


Figure 4. Regeneration frequency of zooid median region (n = 100 animals): (a) migration from the posterior cephalic region to the anterior region of the body; (b) degeneration of the posterior cephalic structures and formation of a new anterior cephalic region; (c) formation of a new anterior cephalic region and body region behind the posterior cephalic region, followed by paratomy; (d) worms died without showing any regenerative pattern.

did not undergo regeneration. In these cases, the ciliated pits and cephalic structures of the zooids were well formed; the worms swam for around one day without showing signs of regeneration and died (Figures 2 and Figures and 3(d)). The frequencies in which these regeneration patterns occurred were: in 14% migration from the posterior cephalic region to the anterior region of the body (Figure 4); in 56% degeneration of the posterior cephalic structures and formation of a new anterior cephalic region; in 13% formation of a new anterior cephalic region and body region behind the posterior cephalic region, followed by paratomy; and in 15% the worms died without showing any regenerative pattern. The differences among regeneration frequencies are significant (chi-squared = 43.24; 3 degrees of freedom; $P < 0.005$).

For the second section scheme, the anterior region of sectioned worms always regenerated (Figure 2(II g)), while the posterior body part never regenerated (Figure 2(II h)). For the median body portion an outcome similar to described to median portion of the first section scheme was observed (Figure 2I (b–e)).

Strong GFP expression in ciliated pits was used to follow the regeneration of fluorescent worms. As seen in Figure 5, in the posterior region containing

part of the intestinal region, soon after sectioning (2 h) the ciliated pit was not observed (Figure 5(a1)). However, 30 h later a fluorescent ciliated pit can be observed, and the worms had completely regenerated (Figure 5(a2)). The median body part showed several regenerative patterns, similar to previously described. 1) New cephalic structures can develop in the anterior part; and the cephalic structure that was on the posterior body part, previous to fission, complete development, forming a worm with two zooids (B2); 2) new cephalic structures are formed in the anterior region and resorption of the posterior cephalic structures occurs (B3); 3) migration of cephalic structures occurs from the posterior region to the anterior portion of the zooid (B4); and 4) new cephalic structures are formed in the anterior region, but regeneration is not complete and the worm dies (B5).

Discussion

The contradictory results described in the literature about the regeneration of *S. leucops* have been attributed, at least partially, to problems in species identification, as suggested by Van Cleave (1929) and

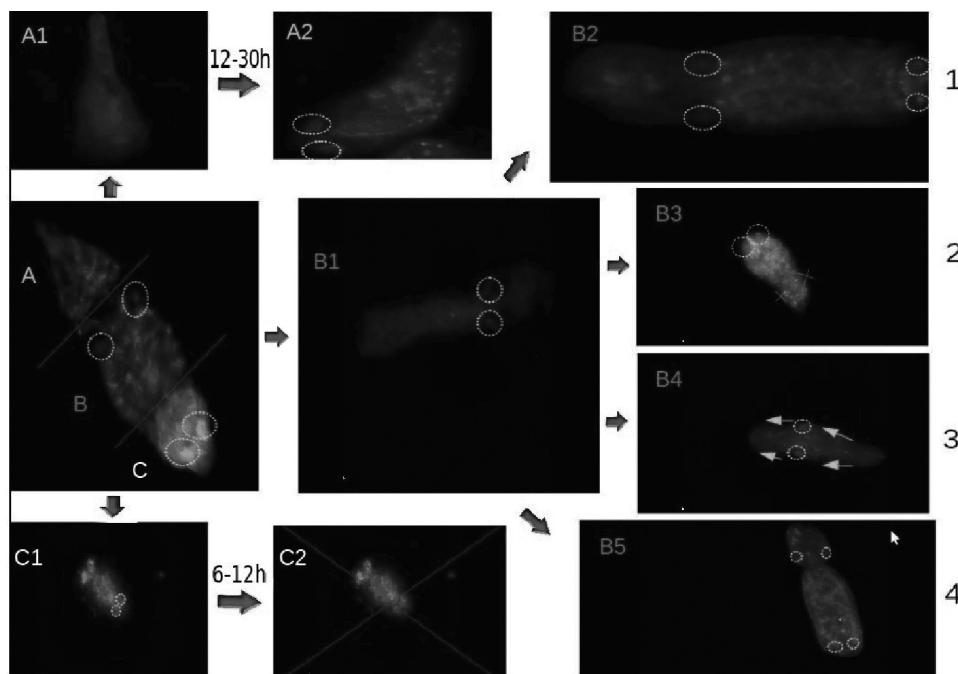


Figure 5. Fluorescent ciliated pits in the regeneration process of *Stenostomum leucops*. Worms of two zooid phases are fissioned in three parts: A-B-C. The fluorescent ciliated pits are indicated by circles. The posterior body part (a) soon (2 h) after fission does not have fluorescent ciliated pits (a1); Completely regenerated (a2), 12 to 30 h after fission show fluorescence in the ciliated pit. The median portion (b) soon after fission (b1) can develop new structures in the anterior part and in the posterior body part, forming a worm with two zooids (b2); a new head is formed in the anterior region and resorption of the posterior head structures occurs (b3); migration of ciliated pits occurs in the anterior region (b4); and new ciliated pits are formed in the anterior region, but regeneration is not complete and the worm dies (b5).

Nuttycombe and Waters (1938). For example, Van Cleave (1929) and Child (1900) described that the specimens they used had five zooids, which does not agree with the described characteristic for this species of having only two zooids. In our experiments using a clonal strain, all patterns of regeneration were as described by Ritter and Congdon (1900), Child (1900, 1902), Hartmann (1922), Ruhl (1927) and Van Cleave (1929). The regeneration process depended on the part of the organism and the stage of development of the zooids. To the anterior and posterior ends, it was fundamental for the sectioned part to contain a small portion of the endodermal digestive sac for regeneration to occur. This was previously pointed out by Child (1900). Van Cleave (1929) suggested that the stage of zoid development is important for regeneration. The GFP marker, expressed in the ciliated pits, allowed us to show that in early stages, cephalic structures can migrate from the posterior to the anterior region of the zooid. In the latter stages of regeneration, when the ciliated pits expressed GFP more intensely, resorption of the posterior cephalic structures and development of new ones in the anterior region typically occur.

Planarians are the principal animal model in research concerning the molecular processes that guide regenerative phenomena. At this time, understanding these processes has gained special importance to support stem cell research, which promises valuable therapies in the biomedical sciences. Many genes have been found in transcriptome research associated with various regenerative processes in planarians (Sandmann et al. 2011; Onal et al. 2012; Resch et al. 2012). As demonstrated in this study, *S. leucops* has a complex pattern of regeneration that is dependent on the body part involved. A transcriptome comparative approach, using previous knowledge obtained in other planarians, may shed light on evolutionary aspects of the regenerative process.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by research grants and fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico—CNPq [421255/2018-8 and 305613/2020-0], Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

ORCID

Elgion L. S. Loreto  <http://orcid.org/0000-0002-7586-8168>

References

- Borkott H. 1970. Geschlechtliche Organisation, Fortpflanzungsverhalten und Ursachen der sexuellen Vermehrung von *Stenopus sthenus* nov. spec. (Turbellaria, Catenulida). *Z Morphol Ökol Tiere*. 67:183–262.
- Campos A, Cummings MP, Reyes JL, Lacleste JP. 1998. Phylogenetic relationships of plathyhelminthes based on 18S ribosomal gene sequences. *Mol Phylogenet Evol*. 10 (1):1–10. doi:10.1006/mpev.1997.0483. PMID: 9751913.
- Child CM. 1900. Fission and regulation. in *Stenostomum leucops*. Meeting of the American Morphological Society. Abstracts. *Biolog Bull* 2(6):329–365.
- Child CM. 1902. Studies on regulation. *Wilhelm Roux' Arch Entwickl mech Org*. 15:187–234.
- Dugès A. 1828. Recherches sur l'organisation et les moeurs des Planariés. *Ann Sci Nat Paris Sér*. 1(15):139–142.
- Gehring WJ. 2002. The genetic control of eye development and its implications for the evolution of the various eye-types. *Int J Dev Biol*. 46(1):65–73.
- Hartmann M. 1922. Über den dauernden Ersatz der ungeschlechtlichen Fortpflanzung durch fortgesetzte Regenerationen. *Bio Zentral*. 42:364–381.
- Horn C, Jaunich B, Wimmer EA. 2000. Highly sensitive, fluorescent transformation marker for *Drosophila* transgenesis. *Dev Genes Evol*. 210(12):623–629. doi:10.1007/s004270000111.
- Hyman LH. 1951. The invertebrates: plathyhelminthes and rhynchocoela the acoelomate bilateria, Vol. II. New York: McGraw-Hill.
- Knakievicz T, Vieira SM, Erdtmann B, Ferreira HB. 2006. Reproduction modes and life cycle of freshwater planarians (Plathyhelminthes, Tricladida, Palucida) from Southern Brazil. *Invertebr. Biol*. 125(3):212–221. doi:10.1111/j.1744-7410.2006.00054.x.
- Larsson K, Ahmadzadeh A, Jondelius U. 2008. DNA taxonomy of Swedish Catenulida (Plathyhelminthes) and a phylogenetic framework for catenulid classification. *Org Divers Evol*. 8 (5):399–412. doi:10.1016/j.ode.2008.09.003.
- Marcus E. 1945. Sobre microturbelarios do Brasil. *Comun Zool Mus Hist Nat Montev*. 1:1–74.
- Noreña C, Damborenea C, Brusa F. 2005. A taxonomic revision of South American species of the genus *Stenostomum* O. Schmidt (Plathyhelminthes: catenulida) based on morphological characters. *Zool J Linn Soc*. 144(1):37–58. doi:10.1111/j.1096-3642.2005.00157.x.
- Nuttycombe J, Waters A. 1938. The American species of the genus *Stenostomum*. *Proc Am Philos Soc*. 79:213–301.
- Onal P, Grün D, Adamidi C, Rybak A, Solana J, Mastrobuoni G, Wang Y, Rahn H-P, Chen W, Kempa S, et al. 2012. Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J*. 31 (12):2755–2769. doi:10.1038/emboj.2012.110.
- Oviedo NJ, Nicolas CL, Adams DS, Levin M. 2008. Planarians: a versatile and powerful model system for molecular studies of regeneration, adult stem cell regulation, aging, and behavior. *CSH Protoc*. pdb.emo101. doi:10.1101/pdb.emo101

- Resch AM, Palakodeti D, Lu Y-C, Horowitz M, Graveley BR. 2012. Transcriptome analysis reveals strain-specific and conserved stemness genes in *Schmidtea mediterranea*. PLoS ONE. 7(4): e34447. doi:10.1371/journal.pone.0034447.
- Rieger RM. 1986. Asexual reproduction and the turbellarian archetype. Hydrobiologia. 132(1):35–45. doi:10.1007/BF00046226.
- Ritter WE, Congdon EM. 1900. On the inhibition by artificial section of the normal fission plane in *Stenostoma*. Proc Calif Acad Sci. 2(6):365.
- Rohde K, Hefford C, Ellis JT, Baverstock PR, Johnson AM, Watson NA, Dittman S. 1993. Contributions to the phylogeny of platyhelminthes based on partial sequencing of 18S ribosomal DNA. International Journal for Parasitology. 23(6):705–724. doi:10.1016/0020-7519(93)90067-9.
- Rosa MT, Pereira CM, Ragagnin GT, Loreto ELS. 2015. *Stenostomum leucops* Dugès, 1828 (PLATYHELMINTHES, CATENULIDA): a PUTATIVE SPECIES COMPLEX WITH PHENOTYPIC PLASTICITY. Pap Avulsos Zool. 55(1):375–383. doi:10.1038/s41598-019-55659-w.
- Rosa MT, Loreto ELS. 2019. The Catenulida flatworm can express genes from its microbiome or from the DNA it ingests. Sci Rep. 9(1):19045. doi:10.1038/s41598-019-55659-w.
- Ruhl L. 1927. Double structure and other anomalies in *Stenostomum*. Zool Anz. 8:180–190.
- Sánchez Alvarado A. 2012. Q&A: what is regeneration, and why look to planarians for answers? BMC Biol. 10(1):88. doi:10.1186/1741-7007-10-88.
- Sandmann T, Vogg MC, Owlarn S, Boutros M, Bartscherer K. 2011. The head-regeneration transcriptome of the planarian *Schmidtea mediterranea*. Genome Biol. 12(8):R76. doi:10.1186/gb-2011-12-8-r76.
- Van Cleave CD. 1929. An experimental study of fission and reconstitution in *Stenostomum*. Physiological Zoology. 2(1):18–58. doi:10.1086/physzool.2.1.30151062.
- Yamazaki M, Asakawa S, Murase J, Kimura M. 2012. Phylogenetic diversity of microturbellarians in Japanese rice paddy fields, with special attention to the genus *Stenostomum*. Soil Sci Plant Nutr. 58(1):11–23. doi:10.1080/00380768.2012.658350.

Capítulo 4

The Catenulida flatworm can express genes from its microbiome or from the DNA it ingests.

OPEN

The Catenulida flatworm can express genes from its microbiome or from the DNA it ingests

Marcos Trindade Rosa¹ & Elgion L. S. Loreto^{2*}

Stenostomum are tiny planarians of the phylum Platyhelminthes that reproduce asexually. We transfected these worms using plasmids containing a *gfp* reporter gene. Here we show that they can express genes present in plasmids carried by bacteria and those that are encoded by naked DNA, such as plasmids or PCR fragments, transfected by electroporation; they can also express genes taken up during feeding. The microbiome associated with worm maintenance was evaluated, and the results indicated that when a plasmid is maintained in the microbiome, *gfp* gene expression is stable. When genes originate from naked DNA or bacteria not maintained in the microbiome, GFP expression is transient. Therefore, changes in the microbiome can modify the ability of worms to express foreign genes. In stable GFP-expressing worms, NSG showed that the *gfp* gene was maintained in the plasmid and was not integrated into the chromosome. These results suggest that, at least for some organisms such as flatworms, the expression of genes provided by the microbiome or the environment can be considered among the potential sources of phenotypic plasticity, which can have implications for evolvability.

The holobiont theory posits that multicellular eukaryotes are not autonomous organisms but rather biological units comprised of the organism and numerous microbial symbionts (holobiont). This postulation brings with it a current debate in evolutionary research regarding the role of the microbiome in evolution^{1–3}. The hologenome consists of the nuclear genome, organelles and microbiome, and it is important for evolution because phenotypic variation can result from the interaction of products of more than one genome in a holobiont³.

A wide spectrum of interactions can be found in different holobionts. In some cases, the microbiome associated with organisms is a source of different materials such as nutrients, probiotics and secondary metabolites or acts in host defense, for example by degrading xenobiotics^{4–6}. In others, the members of a holobiont are completely dependent on one other and, in some organisms, the microorganisms become intracellular components of the host. In several platyhelminths, annelids and molluscs, symbiotic bacteria harboured in the trophosomes are responsible for supplying energy to the holobiont⁷. The *Paracatenula* flatworm (Catenulida), which is devoid of a mouth and gut, is an excellent example of this extreme holobiont interaction. Chemoautotrophic bacteria of the genus *Candidatus* Riegeria are a fundamental symbiont in *Paracatenula*'s trophosomes, corresponding to half of the biomass of these worms and being primary energy source⁸. Organisms with simple body plans, having almost all of their cells in direct contact with the surrounding environment, can have a stronger interaction with the microbiome. In sponges, for example, microbial symbionts comprise up to 35% of the holobiont biomass and are fundamental providers of metabolites for sponge physiology⁹.

In this study we present evidence that the flatworm *Stenostomum*, which is also formed by many cells in direct contact with the environment, shows a host-microbe interaction not described thus far. Their cells can express genes present in plasmids carried by their native bacteria and even genes from naked DNA, such as plasmids or PCR amplicons, taken up during feeding. The flatworm *Stenostomum* belongs to the class Catenulida, a basal group of Platyhelminthes^{10–12}. They reproduce asexually by paratomy, a remarkable process in which the formation of anterior structures in the median region of the body forms zooids. After complete maturation of zooids, each one detaches to form new organisms. In *S. leucops*, the paratomy process occurs in approximately 42 h at 28 °C, producing zooids containing approximately 2400 cells. These flatworms range from 0.5 to 2 mm in length¹³. The first requisite of genetic transformation, the incorporation of foreign DNA into a cell, is the entrance of DNA

¹PPG Biodiversidade Animal, Univ. Fed. de Santa Maria (UFSM), Santa Maria, Brazil. ²Department of Biochemistry and Molecular Biology, CCNE, Univ. Fed. de Santa Maria, Av. Roraima 1000, 97105-900, Santa Maria, RS, Brazil. *email: elgion@base.ufsm.br

into the cell. Normally, eukaryotic cells do not easily accept foreign DNA, and several procedures have been used to promote this step, such as some chemical-based methods (calcium phosphate or liposomes), physical-based methods (electroporation, microinjection, particle bombardment) and viral methods^{14,15}. Transfection has been described for several Platyhelminthes species using particle bombardment or electroporation of RNA or plasmids¹⁶. Nucleic acid entrance into the cell can lead to biological consequences, such as expression of carried ORFs. This gene expression can be stable when the foreign DNA is inserted into the genome or maintained as stable episomes. Expression can also be transient, whereby the maintenance of DNA or RNA in the cells is temporary, and the expression is lost over time¹⁷.

Stable genetic transformation of Platyhelminthes was obtained using a transposon-based vector in planarians¹⁸ and *Schistosoma*¹⁹. This methodology has been widely used in other organisms, in which genes of interest are inserted between the transposon terminal inverted repeats (TIRs) in a vector plasmid. The TIR sequences are recognized by a transposase enzyme encoded by a helper plasmid. Cells transfected with both plasmids can have the sequences of the vector plasmid transposed into their chromosomes²⁰.

The intake of genetic material can promote a transient gene expression, observed when the gene of interest is not integrated into the host genome and is, for instance, maintained as stable episomes. For example, it has been demonstrated that the phenotypes of various invertebrates, when exposed to dietary material containing *in vitro* synthesized dsRNA, or to organisms such as bacteria, artificially expressing dsRNA, are affected by gene silencing through the RNAi process^{21,22}. In these cases, a transient gene silencing occurs while the dsRNA is present into the de host cells, or the bacteria expressing dsRNA are present in the host diet.

In this study, we show that *Stenostomum* can transiently express the *gfp* gene from plasmids contained in bacteria provided in their diet or even from naked DNA such as plasmids or PCR fragments. Also, stable expression of the *gfp* gene was obtained followed by plasmid maintenance in the worms, by some symbiotic bacteria. The evolutionary implications of these findings are discussed in light of the hologenome theory.

Results

Our initial goal was to promote genetic transformation of worms using a transposon vector and *gfp* as a reporter gene, as described for planarians by González-Estévez *et al.*¹⁸. *S. leucops* of the SL0-sm01 and SL0-sm02 strains transfected by electroporation with pBac[3xp3-EGFPafm] and pBΔSac plasmids had survival rates of 20+/- 10% for both strains. However, GFP expression differed between tested strains (Fig. 1). For SL0-sm01, 80+/- 5% of worms showed strong GFP expression 24 hours after electroporation, reaching 100% in 96 hours. Transfected worms of this strain showed stable GFP expression, transmitted to subsequent generations for six months (see Supplementary Movie 1). The SL0-sm02 strains showed weaker and transient GFP expression. Twenty-four hours after electroporation, 80+/- 5% of worms showed GFP fluorescence. However, GFP expression decreased and was lost after 96 hours.

Aiming to verify if stable GFP expression observed in SL0-sm01 strain was a result of transposition of this gene into the *S. leucops* genome or the result of a plasmid being maintained in the worm, PCR-based assays were performed three and fifteen days after electroporation (Fig. 2, Supplementary Table 4). In SL0-sm01 worms, amplicons containing regions outside of the TIR sequences of the transposon vector were obtained. This suggests that, if integration into the *S. leucops* chromosomes had occurred, it was not a result of precise excision mediated by the *piggyBac* transposase (Fig. 2A,C; Supp. Fig. 1). Alternatively, the complete plasmids could be maintained in the worms. For SL0-sm02, amplicons were obtained three days after electroporation but not fifteen days after, showing that the transfected plasmids had been lost (Fig. 2B). These results are concordant with the transient GFP expression observed in SL0-sm02 strain.

To distinguish between the hypotheses of sequence integration into chromosomes and plasmid maintenance in the SL0-sm01 strain, the genome of these worms was sequenced using an NGS approach. The complete sequence of plasmids pBac[3xp3-EGFPafm] and pBΔSac were assembled with a high coverage (58X). Approximately 2,000 reads containing sequences of the pBac[3xp3-EGFPafm] plasmid were obtained, which allowed the assemblage of the complete plasmid, including the *gfp* gene. However, among those 2,000 reads, none had “hybrid” sequences (i.e. sequences belonging to both the genome of *Stenostomum* and the transfected plasmid), suggesting that no integration into the worms’ genome occurred. The assembly metrics are presented in Supplementary Table 1. The transformed animals maintained fluorescence for six months, corresponding to around 80 generations after genetic transformation. These animals belonging to the 80 generation were used for genomic DNA isolation for NGS. Therefore, the complete plasmids were maintained in the worm culture during this time.

Aiming to test if the different pattern of plasmid maintenance observed in SL0-sm01 and SL0-sm02 strains could be related to the microbiome associated with the worms, we analysed the microbiomes of these strains using a 16S rRNA gene amplicon NGS approach (Fig. 3). In strain SL0-sm01 four bacterial taxa were found, with *Bacillus cereus* sp. being the most abundant group (85.15%) followed by *Escherichia coli* (14.74%), *B. subtilis* (0.06%) and *B. aeneaus* (0.05%). In strain SL0-sm02 thirty-six bacteria species were found. *E. coli* was absent, and the *Bacillus cereus* species group corresponded to only 1.2% of sequenced reads. These results show that the microbiome is quite different between the strains used. More significantly, these results suggest that the plasmid could be maintained by *E. coli*, as the used plasmid has the ColE1 replication origin, which is related to this bacterium and other closely related species.

To test the hypothesis that plasmids maintained by *E. coli* could be the gene source of GFP expression, we fed the worms of SL0-sm02 strain with *E. coli* X11 (a laboratory strain) containing the plasmids. As can be seen in Fig. 4A, the worms fed with this bacterium expressed GFP. This expression was transient, lasting for three days. The same occurred when worms were fed with “naked DNA” as plasmid or PCR fragments containing the *gfp* gene (Fig. 4B,C), suggesting that GFP expression requires DNA containing the *gfp* gene and not the bacteria. Reinforcing this hypothesis, we showed that *E. coli* X11 harbouring the plasmid with the *gfp* gene does not

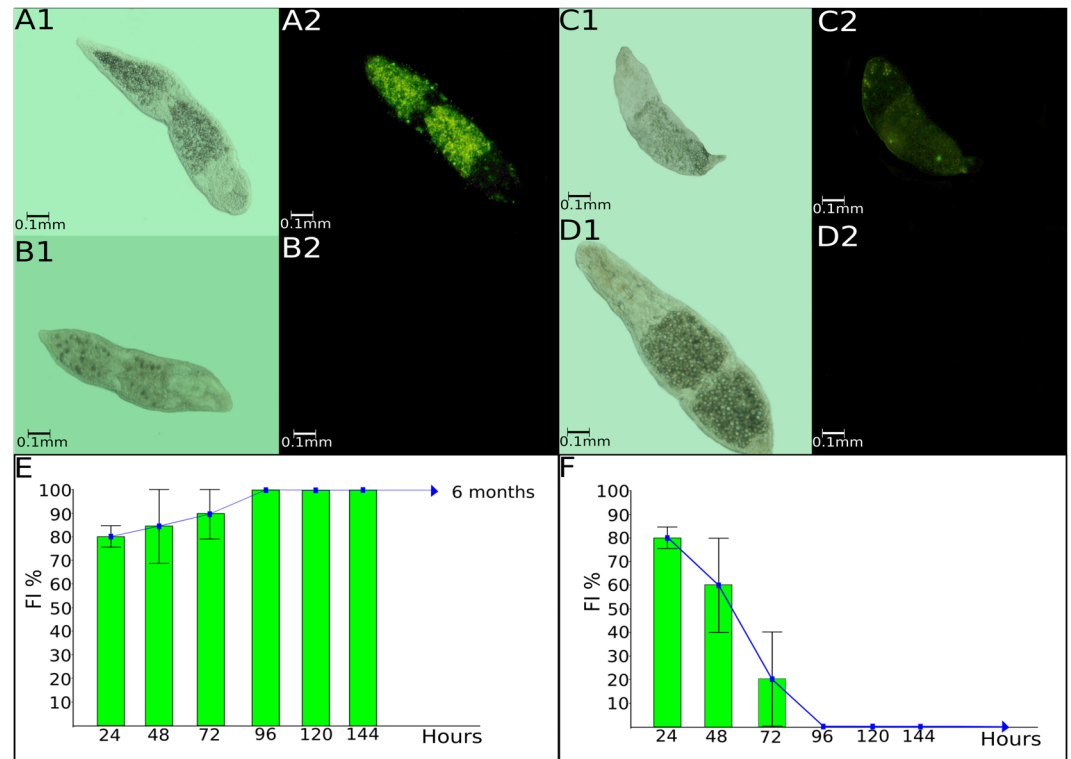


Figure 1. *Stenostomum leucops* transfected with pBac[3xp3-EGFPafm] and p Δ Sac plasmids. Bright-field microscopy of SL0-sm01 worms electroporated with plasmids (A1) and under fluorescence microscopy (A2), showing fluorescence; (B1) SL0-sm01 worm electroporated without plasmids (control) not showing fluorescence (B2); (C1) SL0-sm02 worms electroporated with plasmids and under fluorescence microscopy, showing fluorescence (C2); D1- SL0-sm02 worm electroporated without plasmids (control) not showing fluorescence (D2); (E) percentage of fluorescent individuals (FI) for SL0-sm01 worms electroporated with plasmids, in relation to the time (hours) after electroporation; (F) percentage of fluorescent individuals (FI) for SL0-sm02 worms electroporated with plasmids. Eighty worms were visualized under the microscope in each assay.

fluoresce (Fig. 4D). Instead, by disintegrating the worms, allowing isolated cells to be seen, we could observe fluorescence in the cytoplasm of cells from worms fed with plasmids containing the *gfp* gene but not in the cells of worms that did not receive the plasmids (Fig. 4E,F). These results point out that the *gfp* gene contained in DNA present in the *Stenostomum*'s feed can be expressed in the worm's cells. In conclusion, the transfected *gfp* gene was shown to be transiently expressed by the worm when presented as “naked DNA” or in plasmids carried by bacterial strains that cannot be maintained as part of *Stenostomum* microbiome (such as the auxotrophic laboratory strain of *E. coli* XL1). However, when the plasmid is harbored by native *E. coli*, the GFP expression becomes stable over time.

A remarkable observation was the pattern of GFP expression in *S. leucops*, in both strains tested, when the pBac [3xp3-EGFPafm] plasmid was used. The GFP expression was spread across the whole body of the worm, and more intense expression was observed in the ciliated pits (Fig. 5A). The ciliated pits are chemosensory structures, comparable to the eyespots of other planarians. The *gfp* gene used in our study is under control of a promoter with the Pax6 dimeric binding sites, and it is specifically expressed in the eye tissue of various arthropod and planarian eyespots. Although variable in frequency, GFP expression in the ciliated pits was observed in worms fed bacteria or “naked” DNA harbouring *gfp* gene (Fig. 5C–E). We decided to test the reporter gene under other regulatory sequences present in different plasmids. As it is demonstrated in Fig. 5B, worms harboring bacteria carrying the *mCherry* gene, under control of the Histone H3 promoter, displayed red fluorescence spread over most of their body.

A low-coverage transcriptome was assembled from worms fed with pBac [3xp3-EGFPafm] plasmid, aiming to analyse the plasmid sequences that were transcribed in the worms. Only reads covering the *gfp* gene were obtained, suggesting that only this gene was transcribed (Fig. 5F).

To show that *E. coli* harbouring plasmids is able to maintain GFP fluorescence in a stable pattern and, because the laboratory strain (XL1) is auxotrophic and unable to be maintained in worms, we sought *E. coli* strains in nature that could be maintained in worms. Bacterial colonies were isolated on LB plates after electroporation of *Stenostomum* sp. Ssp-sm09, freshly collected from the wild, with the pBac[3xp3-EGFPafm] plasmid. These isolated bacteria can be maintained in a stable pattern in both worm strains, SL0-sm02 and Ssp-sm09, and the worms began to express GFP in a stable pattern, in both strains (Fig. 6A–C; Suppl. Fig. 2). These bacteria were confirmed as *E. coli* by Sanger sequencing of the ribosomal 16S sequence. Also, the microbiome of *Stenostomum* sp. Strain Ssp-sm09 was analysed and shown to be composed of 133 bacterial species, among them *E. coli* (Fig. 6D).

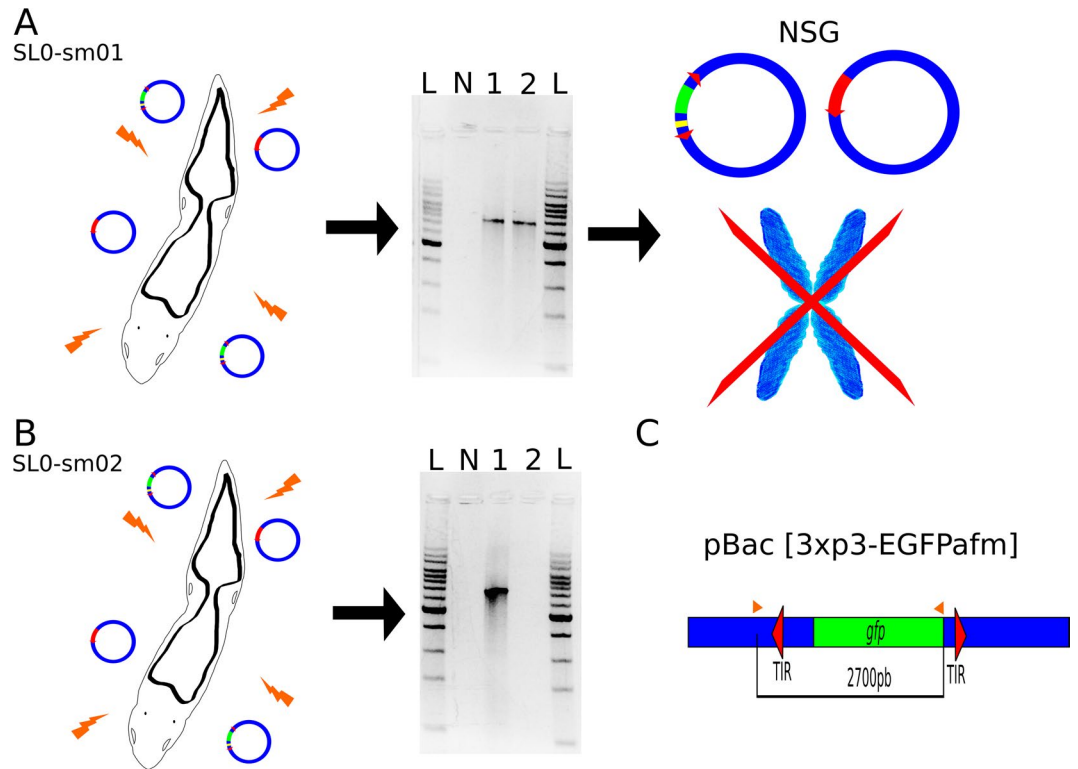


Figure 2. Experimental scheme testing the integration of the *gfp* gene into the *Stenostomum* genome. (A) SL0-sm01 electroporated worms using pBac [3xp3-EGFPafm] and pB Δ Sac plasmids were tested by PCR (gel lanes: L = Kb ladder; N = control/ no DNA; 1 = worm sample three days after electroporation; 2 = worm sample 15 days after electroporation). Amplicons of 2,700 bp indicated that precise transposition had not occurred (see scheme C). Illumina NSG using gDNA isolated six months after electroporation allowed assembly of both plasmids (complete and circularized). However, no indication of integration into the worm's chromosomes was obtained. (B) SL0-sm02 electroporated worms using pBac[3xp3-EGFPafm] and pB Δ Sac plasmids tested by PCR 3 and 15 days after electroporation. No amplicons were obtained 15 days after electroporation, showing the plasmid was lost in this period; (C) scheme of the pBac[3xp3-EGFPafm] plasmid containing *piggyBac* TIRs sequences (red arrowhead) and *gfp* gene (green), other plasmid sequences (blue) and primers used for PCRs (orange arrows). Amplicons containing 81 bp of plasmid sequences upstream of the TIRs, suggest that a precise transposition did not occur.

A comparison of the microbiomes found in the three strains of *Stenostomum* analysed showed that it was highly variable in the number and composition of species. Of a total of 156 species, only one bacterial taxon (*Bacillus cereus* sp. group) was shared by all strains. One species was shared by SL0-sm01 and Ssp-sm09 (*E. coli*), and 11 species were shared by SL0-sm02 and Ssp-sm09 (Supplementary Table 2).

Discussion

Foreign gene expression, even from naked DNA such as a plasmid, is a well-documented phenomenon in eukaryotes. This biological property forms the basis for the promising technology of DNA vaccines, in which DNA contains a gene encoding an antigen, and a promoter/terminator is expressed in mammalian cells, inducing immunity^{23,24}. However, DNA uptake is not a spontaneous process, and several transfection procedures are normally used to induce this step¹⁴. Several studies have shown that, in mammals, meal-derived DNA fragments of up to 300 bp can avoid degradation in the digestive system, enter the circulation system and be found in the blood²⁵. The novelty we introduce with our study, at least to our knowledge, is that the worm *Stenostomum* can spontaneously take up foreign DNA provided by different routes, including by feeding, and can express the protein-coding sequences that it contains.

Transient gene expression induced by feeding on genetic material was first shown in *C. elegans* eating double-stranded RNA (dsRNA) or bacteria expressing dsRNA. These feeding materials have the ability to silence genes by dsRNA-mediated genetic interference (RNAi)²¹. Dietary exposure of several invertebrate organisms has shown similar results²². In Platyhelminthes, feeding the freshwater planarian *Schmidtea mediterranea* on *E. coli* expressing double-stranded RNA inhibits gene expression by RNAi²⁶. These dsRNA involved in feeding-mediated RNAi are molecules <200 nucleotides in length that typically do not code for a protein. In *C. elegans*, dsRNA uptake is mediated by a channel protein (SID-1)²⁷. Homologues to this protein are found in other organisms although, in several insects, SID-like proteins are not involved in dsRNA uptake. In *Drosophila*, the entrance of dsRNA into the cell involves clathrin-mediated endocytosis²⁸. In summary, these studies have shown that diet-derived dsRNA is able to modulate gene expression by RNAi. However, in the present work, a different

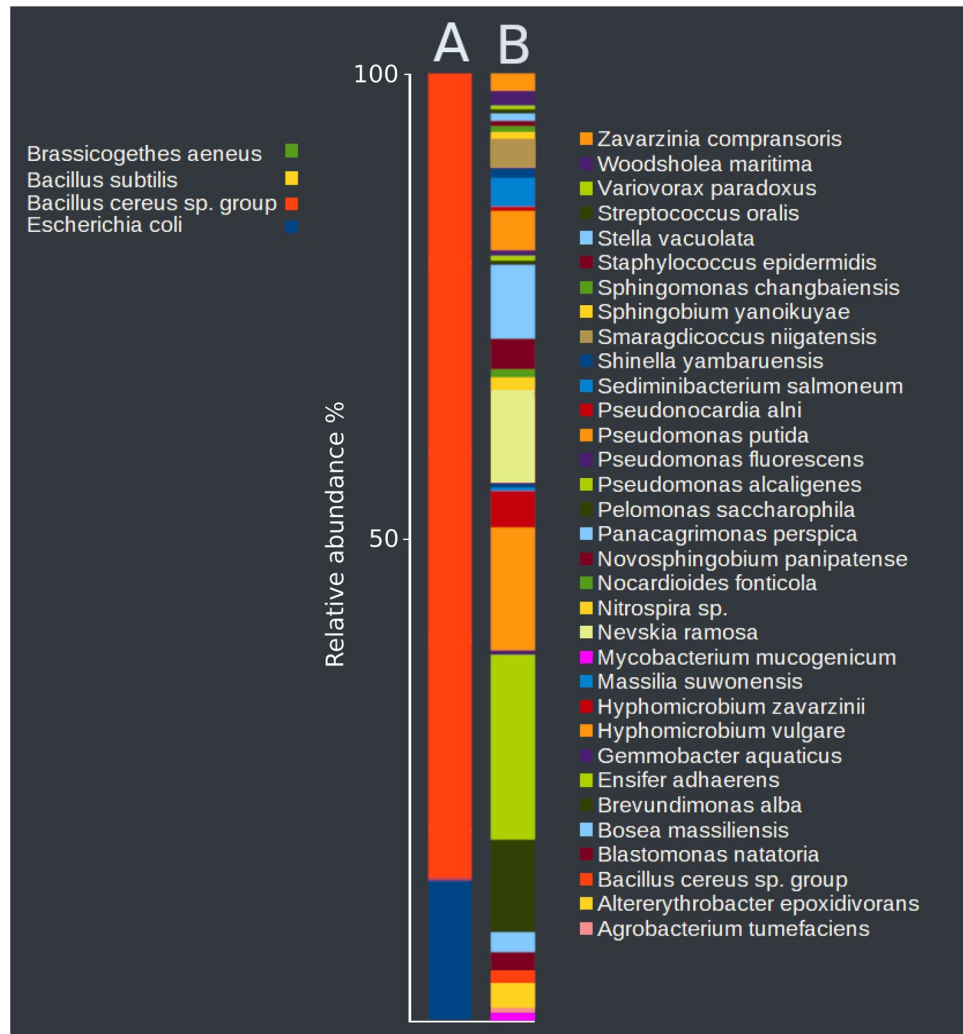


Figure 3. Microbiome associated with *Stenostomum leucops*: (A) microbiome present in the SL0-sm01 strain; (B) microbiome present in the SL0-sm02 strain. The bar plots represent relative abundance (in %).

phenomenon was described, in which phenotypic expression is related to ingested DNA and not dsRNA. Also, the sizes of the fragments involved were larger than those described for dsRNA uptake, since the smallest DNA fragment we used was 3394 bp, the largest plasmids are approximately 9 kbp and the CDS contained in the ingested DNA was expressed.

The expression of foreign genes is normally transient, mainly in mitotically active cells, when organisms or cells are transfected with plasmids, DNA fragments or episomes¹⁷. However, more stable gene expression from plasmids has been described in the literature. For example, plasmids persisted in mouse skeletal muscle for 19 months, expressing the *luciferase* gene after intramuscular injection²⁹. Similarly, we observed stable GFP expression in transfected *Stenostomum* worms, when they hold up in their microbiota. However, a difference can be highlighted: rodent muscle cells are post-mitotic, meaning they do not divide further, which can explain plasmid maintenance in these cells²⁹. In contrast, *Stenostomum* is an organism that undergoes intense cell division, and stable expression was associated with the maintenance of *E. coli*. The plasmids used in the transfection had a replication origin for this species (ColE1), meaning that they can be maintained in *E. coli* and related species. The absence of *E. coli* in strain SL0-sm02 could explain its inability to stably express GFP.

The fluorescence observed in the worms could be from GFP expressed by the ingested bacteria. However, some evidence indicates that DNA uptake by worm cells is the more plausible explanation: (i) bacterial colonies grown in Petri dishes did not show fluorescence; (ii) naked DNA containing the *gfp* gene is enough to produce fluorescence; (iii) the artificial Pax-6 promoter is recognized by the transcriptional machinery of eukaryotes, but not by bacteria, and there was marked expression of GFP in the ciliated pit; (iv) fluorescence is observed in the cytoplasm of worms cells; (v) transcriptome analyses showed only transcripts of the *gfp* gene, suggesting an eukaryotic transcription pattern.

An increasing number of examples have been reported showing that some phenotypes are not the result of only one genotype but rather, the interaction of products provided by two or more genotypes, i.e. the hologenome. In some cases, remarkable relationships have evolved; for example, the Hawaiian bobtail squid has a light organ to

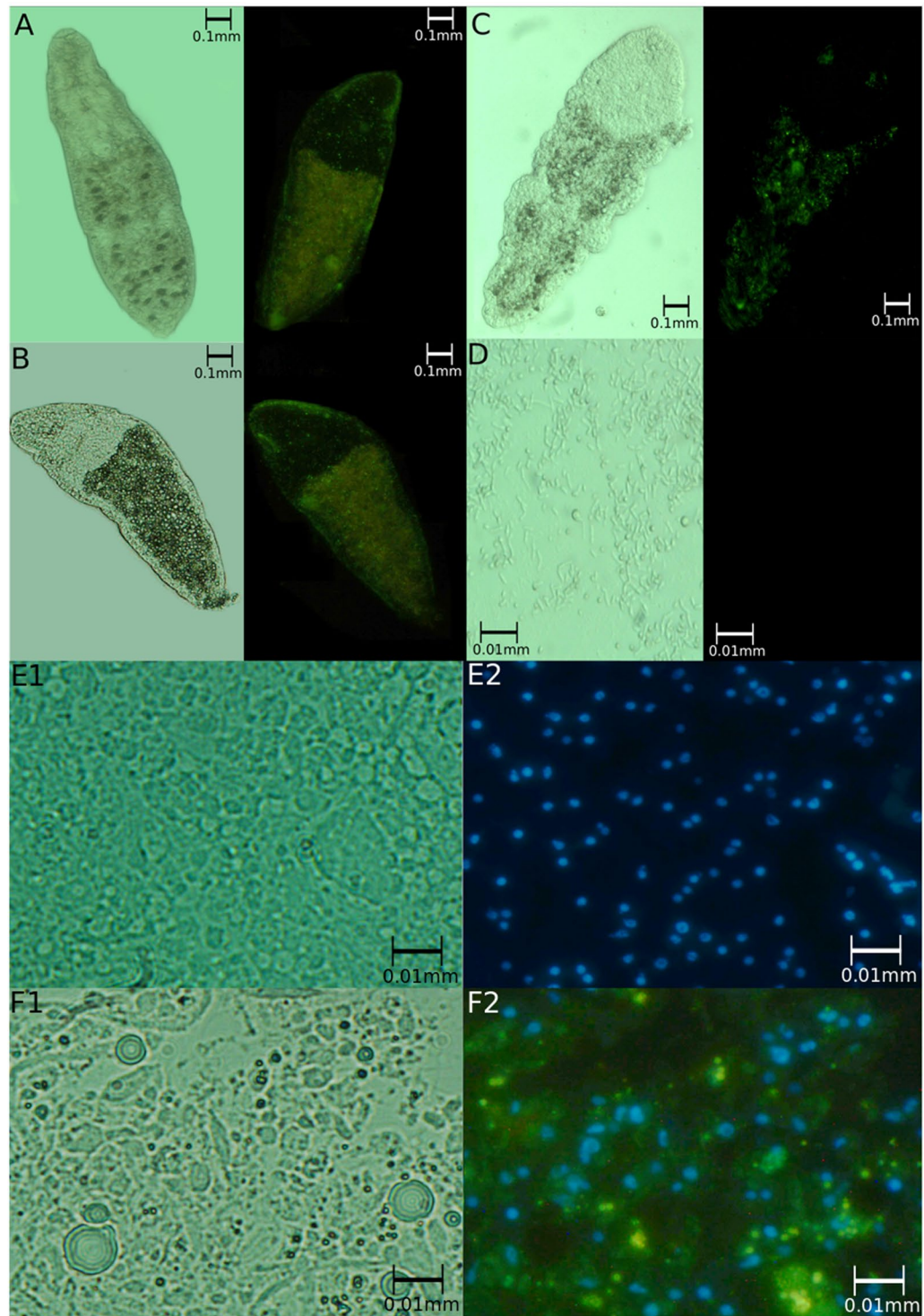


Figure 4. GFP expression in worms fed (A) *E. coli* XL1 containing pBac [3xp3-EGFPafm] plasmid; (B) only the pBac [3xp3-EGFPafm] plasmid; (C) 3.2kb PCR fragment containing complete *gfp* gene. (D) the bacteria *E. coli* XL1 harbouring the pBac [3xp3-EGFPafm] plasmid; E1) *Stenostomum* cell, stained with DAPI, in bright-field microscopy; E2) in fluorescence microscopy (blue = nucleus); F1) *Stenostomum* cell of worms fed *E. coli* XL1 containing pBac [3xp3-EGFPafm] plasmid, stained with DAPI, in bright-field microscopy; F2) in fluorescence microscopy (blue = nucleus; green = GFP). Sixty worms were visualized under the microscope in each assay.

harbour fluorescent bacteria and, using an extraordinary mechanism of molecular signalling between the squid and the endosymbiont, the two organisms are able to glow at night in a coordinated manner³⁰. Furthermore, the microbiome is now known to have a significant role in metabolism, immune defense, and behaviour³¹. Symbiosis

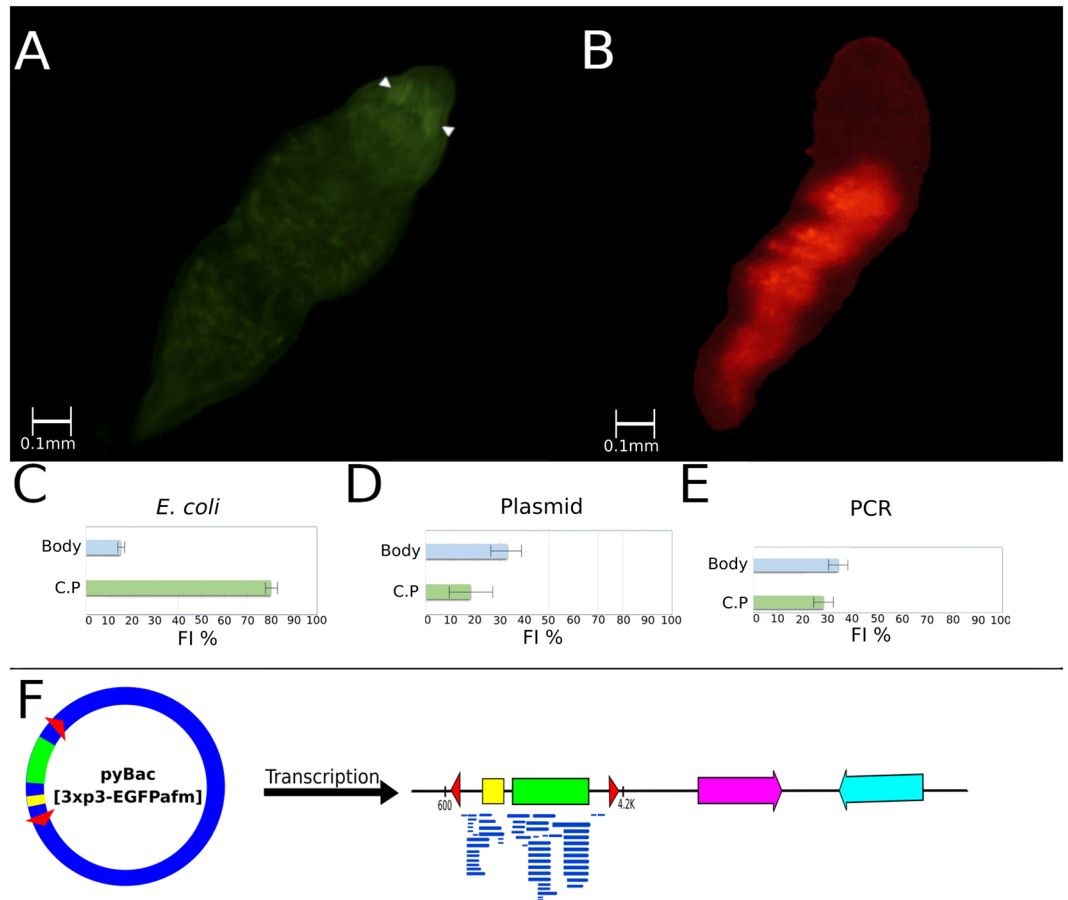


Figure 5. Expression of the *gfp* reporter gene under the control of different promoters. (A) GFP expression mainly in the ciliated pits (CPF, arrowhead) when plasmids have Pax6 regulatory elements in the promoter; (B) pattern of diffuse fluorescence (DF) spread over the body observed in worms expressing mCherry red fluorescent protein under control of the histone H3 promoter; (C) percentage of worms showing diffuse fluorescence spread over the body (blue bars) or in the ciliated pit (green bars), for worms fed with: *E. coli* carrying 3xp3-EGFPafm (with Pax6 RE) plasmid or only the plasmid (D); (E) PCR fragments containing the GFP gene with Pax6 RE (green bars). Whiskers represent the standard deviation. (F) Transcriptome assay for worms fed 3xp3-EGFPafm plasmid showed reads covering only the *gfp* gene region.

can produce selectable genetic variation for the holobiont, and it can generate reproductive isolation, which has significant evolutionary implications³². In sponges, which are organisms with a simple body plan, the cells are in direct contact with the surrounding medium, and the bacterial symbionts are of central importance for their metabolism and are major contributors to the evolutionary and ecological success of these organisms⁹. Ctenulida worms, like sponges, have a simple body plan and cells in direct contact with the surrounding medium, which may be involved in their capacity to obtain genetic material from the environment and express it. As these materials can be provided by the microbiome, this provides a new route for holobionts to increase their arsenal of evolutionary tools. Ctenulida show extraordinary phenotypic plasticity¹³. Also, their microbiome, as shown here, can be highly variable. It is worth mentioning that, in our study, only the eubacteria representative of the microbiome were characterized. Thus, the microbiome probably could be more diverse if the eukaryotic and archaeal counterparts were also added to the analysis. The possibility that worms can express genes harboured by their microbiome may contribute to their phenotypic plasticity and, therefore, to their evolvability.

Methods

Animals used and their maintenance. Two different *Stenostomum leucops* strains were used in this work, SL0-sm01 and SL0-sm02. They were collected in a pond at Santa Maria, Brazil (53°17'W; 29°28'S) in 2009 and maintained in the laboratory at 28 °C, in reconstituted water and fed with powdered milk Nestle/Molico¹³. Both strains had the same clonal origin. However, after several years of culture, the original strain showed a reduction in the reproduction rate. To increase the culture vigour we collected water from a dam, centrifuged it at 112 × g for 5 minutes, and added the supernatant to the culture, aiming to improve the worms' microbiota. The centrifugation step aimed to prevent other worms from contaminating the culture. The original strain came to be called SL0-sm01 and the strain which received the bacteria called SL0-sm02.

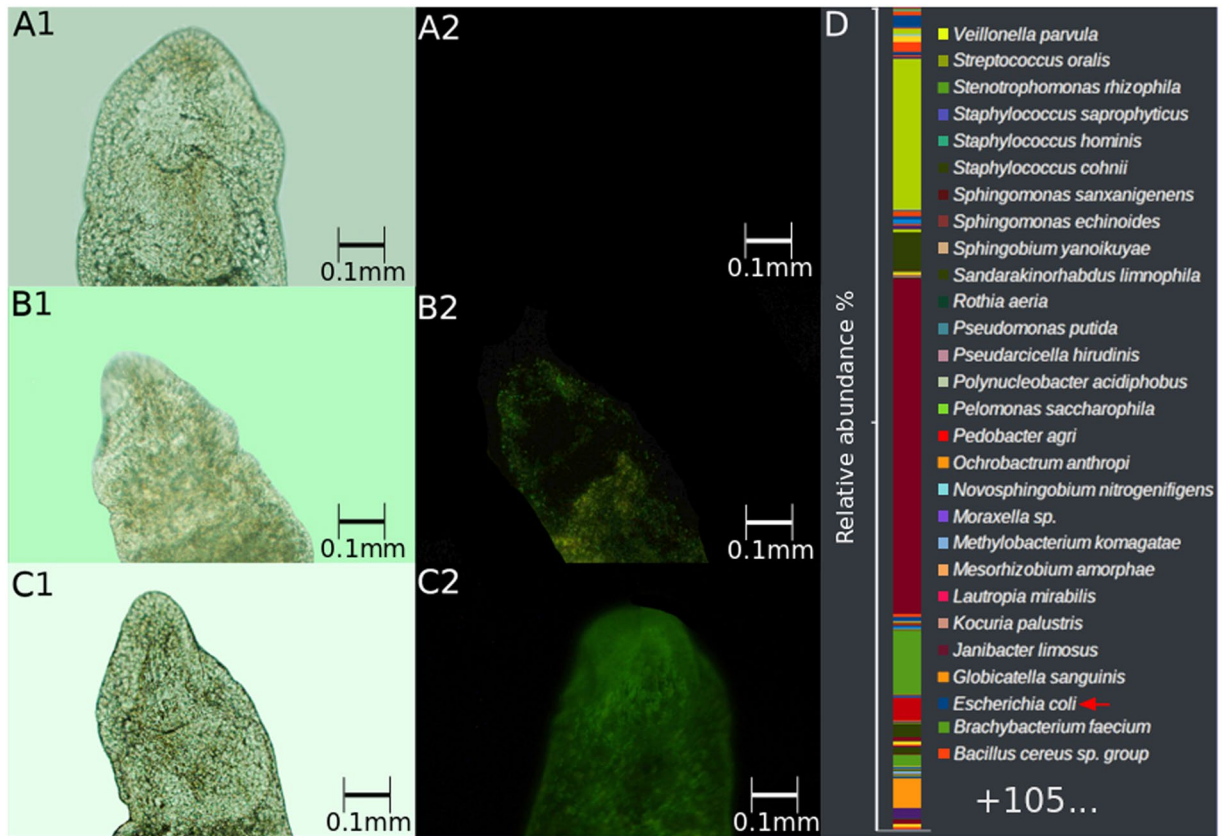


Figure 6. Stable GFP expression induced by the new *E. coli* strain isolated in nature, harbouring the pBac [3xp3-EGFPafm] plasmid: (A) Ssp-sm09, control - worms fed *E. coli* without plasmid (A1 = bright-field; A2 = fluorescence microscopy); (B) Ssp-sm09 fed with *E. coli* harbouring pBac[3xp3-EGFPafm] plasmid (B1 = bright-field; B2 = fluorescence microscopy); (C) Ssp-sm09 fed “naked” pBac[3xp3-EGFPafm] plasmid (C1 = bright-field; C2 = fluorescence microscopy); (D) microbiome present in the Ssp-sm09 strain. The bar plots represent relative abundance (in %).

A third strain, Ssp-sm09, which belongs to an undescribed *Stenostomum* species, was collected in a pond at Santa Maria, Brazil (53°17'W; 29°28'S) in 2019 and maintained in the laboratory under the same conditions used to maintain the other strains. DNA barcoding (accession number MN257507) showed that this species is closely related to *S. sthenum*.

Plasmids and PCR fragments used. Three different plasmids were used for *S. leucops* transfection: (i) pBac [3xp3-EGFPafm]³³ containing the *piggyBac* transposon Terminal Inverted Repeats (TIRs) and the green fluorescent protein ORF under the control of an artificial promoter containing three Pax-6 homo-dimer binding sites. This promoter was constructed to drive strong expression of the GFP protein in the eye tissue of *Drosophila* and to promote expression of this gene in other organisms, as in planarian eyespots¹⁸; (ii) pΔSac – helper³⁴, which contains the *piggyBac* transposase gene, but one of its TIRs was inverted, making it unrecognizable to transposase, producing a non-mobilizable transposable element copy. However, this plasmid promotes the production of *piggyBac* transposase to mobilize sequences present in the pBac [3xp3-EGFPafm] plasmid; (iii) pjaf15_H3p-Kat-H3T plasmid containing the mCherry reporter gene under control of the *Cryptococcus neoformans* Histone H3 promoter. As a control, worms were fed *E. coli* without plasmids or with plasmids not containing *gfp* reporter genes (pCR 2.1, Invitrogen; Carlsbad, Ca, USA).

A PCR product was used for transfection, corresponding to the amplicon of 3,394 bp obtained from pBac [3xp3-EGFPafm] containing the *gfp* ORF, the 3PX3 promoter region and the *piggyBac* TIRs. Primers and PCR conditions can be found in Suppl Table 3 and Suppl. Fig. 1.

Transfection procedures. Three different procedures were used for transfection of *Stenostomum*: (i) electroporation; (ii) feeding animals with *E. coli* XL1 harbouring the desirable plasmid or (iii) feeding or exposing the animal to “naked DNA” (isolated plasmid or PCR products).

During the electroporation process, the worms were exposed to a single pulse of 80 V for 1 ms, using a Gene Pulser Xcell Electroporation System (Bio-Rad, Hercules, Ca, USA) apparatus. Fifty worms were put in a 2-mm cuvette with 200 µl of reconstituted water, 35 µl of PBS and 1.5 µg of DNA. After electroporation, the worms were rinsed in reconstituted water and maintained individually in wells of a plastic cell culture plate with reconstituted water (3 ml per well). For each assay, eight independent experiments were performed.

In the second procedure, the worms were fed with *E. coli* XL1 transformed with the plasmids described above. Bacteria were grown in 2 ml of LB medium for 18 hours at 37 °C. Then, the medium was centrifuged and the pellet added to 0.1 mg of powdered milk and maintained for 3 hours at 37 °C in an oven until the mixture was completely dehydrated. After this, the dried mixture was added to 2 ml of reconstituted water, in a well of a cell culture plate containing 10 worms. For each assay, 15 independent experiments were performed.

Native *E. coli* strains were isolated from *Stenostomum sp.* Ssp-sm09, freshly collected from the wild. The worms were electroporated with the pBac[3xp3-EGFPafm] plasmid and, after homogenization, were plated on LB agar medium supplemented with 35 µg/ml of ampicillin. Bacteria obtained by this process were grown in LB medium, dehydrated in powdered milk, as previously described, and used to feed SLO-sm02 and Ssp-sm09 strains.

The third procedure consisted on directly exposing the animals to “naked DNA”. This was done by adding DNA directly into the water in which the animals were maintained (0.7 µg/ml) or adding DNA to the worms’ food. A quantity of 0.7 µg of plasmid DNA or PCR fragments was added to 1 mg of powdered milk. The mixture was incubated in an oven for three hours at 37 °C, until complete dehydration. The dried mixture was added to 2 ml of reconstituted water in a well of a cell culture plate containing 10 worms. For each assay, 15 independent experiments were performed.

The fluorescence of worms was compared across treatments and the control. The worms of this strain showed a very low background fluorescence, differing strongly from the transfected worms. GFP expression was registered using an Olympus BX41 fluorescence microscope (Olympus, Shinjuku, Japan). Pictures were taken using an absorption filter at 475 nm (blue) and an emission filter at 509 nm (green).

For cellular localization of GFP expression, the rostrum region of the worms was stained using DAPI (0,2 µg/ml in PBS) and registered under bright-field and fluorescence microscopy.

DNA extraction and whole genome sequencing. Genomic DNA was isolated from an isoline expressing GFP. This isoline was transformed by electroporation using the plasmids pBac [3xp3-EGFPafm] and pBΔSac – helper. Molecular analysis was performed on the 80th generations after genetic transformation. DNA was isolated from a pool of 100 individuals using a previously described protocol³⁵. A genomic DNA library was constructed from 200 ng of DNA using Illumina TruSeq Nano DNA kit. A MiSeq Reagent Kit v3 (600 cycles) was used for sequencing to produce paired-end reads. Sequencing was performed using an Illumina MiSeq platform by Unidade de Genômica Computacional Darcy Fontoura de Almeida/LNCC/ Brazil (BAM accession number: SAMN12500763).

Raw genomic data were processed to removal of sequencing adapters and low-quality reads using the Trim Galore³⁶ tools. To assemble the target regions of the genome, i.e. the green fluorescent protein gene (*gfp*), the software MITOBIM³⁷ was used with default parameters and using the *gfp* sequence as a seed.

Microbiome analyses. For microbiome analyses, DNA was isolated as previously described. The microbiome analysis was performed by Neoprosecta (<https://neoprosecta.com>). Regions of the 16S ribosomal gene were amplified by PCR and sequenced with an Illumina platform, as described previously^{38,39}. The primers and PCR conditions used are available in Supplementary Table 3. Bioinformatic analysis was performed to implement a blast search using a proprietary databank (Neoprosecta), which contained only 16S sequences of validated taxa³⁸. The microbiome sequencing reads are available at the accession number: SAMN12561444 (*Stenostomum leucops* SLO-sm01), SAMN12561445 (*Stenostomum leucops* SLO-sm02) and SAMN12561446 (*Stenostomum sp.* Ssp-sm09).

The isolated bacterial strains obtained from *Stenostomum sp.* Ssp-sm09, have the 16S ribosomal gene amplified using the same primers used in microbiome analyses (Supplementary Table 3). The PCR product was marked using the BigDye kit and sequenced by using the Sanger method in an AB 3500 platform (Thermo Fischer Scientific, Waltham, Ma, USA).

Transcriptome analyses. Aiming to verify which regions of plasmid are expressed in transfected worms, we performed a low-coverage transcriptome analysis. The *Stenostomum sp.* Ssp-sm09 strain was chosen because this species is larger, and fewer animals are necessary for library preparation. The animals were fed with powdered milk and plasmid as described in the transfection section. Twenty-four hours later, they were checked in a microscope for fluorescence expression, and total RNA was extracted from 30 worms using TRIzol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s instructions. For library preparation, mRNA was enriched with the Dynabeads mRNA Purification kit (Invitrogen), according to the manufacturer’s protocol. Sequencing was performed in an Ion Torrent S5 sequencer (Thermo Fisher Scientific) using the Ion 540 Kit-OT2 and the Ion 540 Chip. Data were collected using Torrent Suite v4.0 software.

The assemblage of reads was performed by the *de novo* approach using Trinity⁴⁰ software available in the Galaxy⁴¹ platform, with default parameters. The assembled contigs were blasted against the plasmid pBac [3xp3-EGFPafm]³³. The reads were mapped using bowtie2⁴² and displayed using Tablet⁴³ (RNA-seq accession number: SAMN12497666).

Received: 2 September 2019; Accepted: 28 November 2019;

Published online: 13 December 2019

References

- Bordenstein, S. R. & Theis, K. R. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biol.* **13**, e1002226 (2015).
- Douglas, A. E. & Werren, J. H. Holes in the Hologenome: Why host-microbe symbioses are not Holobionts. *M Bio.* **7**, e02099 (2016).
- Haag, K. L. Holobionts and their hologenomes: Evolution with mixed modes of inheritance. *Genet Mol Biol.* **41**, 189–197 (2018).

4. Promi, D., Parizad, B. & Jens, N. Metagenomic analysis of microbe-mediated vitamin metabolism in the human gut microbiome. *BMC Genomics* **20**, 208 (2019).
5. Leah, G., Sarah, W. & Libusha, K. The human gut chemical landscape predicts microbe-mediated biotransformation of foods and drugs. *eLife* **8**, e42866 (2019).
6. Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R. & Goodman, A. L. *Nature* **570**, 462–467 (2019).
7. Dubilier, N., Bergin, C. & Lott, C. Symbiotic diversity in marine animals: The art of harnessing chemosynthesis. *Nat Rev Microbiol* **6**, 725–740 (2008).
8. Oliver, J. *et al.* Chemosynthetic symbiont with a drastically reduced genome serves as primary energy storage in the marine flatworm *Paracatenula*. *PNAS* **10**, 1073 (2018).
9. Webster, N. S. & Thomas, T. The sponge hologenome. *MBio* **7**, e00135–16 (2016).
10. Telford, M. J., Lockyer, A. E., Cartwright-Finch, C. & Littlewood, D. T. J. Combined large and small subunit ribosomal RNA phylogenies support a basal position of the acoelomorph flatworms. *Proc. R. Soc. London B* **270**, 1077–1083 (2003).
11. Wallberg, A., Curini-Galletti, M., Ahmadzadeh, A. & Jondelius, U. Dismissal of Acoelomorpha: Acoela and Nemertodermatida are separate early bilaterian clades. *Zool. Scripta* **36**, 509–523 (2007).
12. Rosa, M. T., Oliveira, D. S. & Loreto, E. L. S. Characterization of the first mitochondrial genome of a catenulid flatworm: *Stenostomum leucops* (Platyhelminthes). *J. Zool. Syst. Evol. Res.* **55**, 98–105 (2017).
13. Rosa, M. T., Pereira, C. M., Ragagnin, G. T. & Loreto, E. L. S. *Stenostomum leucops* Duges, 1828 (Platyhelminthes, Catenulida): A putative species complex with phenotypic plasticity. *Pap. Avulsos Zool.* **55**, 375–383 (2015).
14. Evans, T. C. Transformation and microinjection. WormBook, The *C. elegans* Research Community. *WormBook* **1**(108), 1 (2006).
15. Robbins, D. R., Tahara, H. & Ghivizzani, H. C. Viral vectors for gene therapy. *Trends Biotechnol.* **16**, 35–40 (1998).
16. Moguel, B., Bobes, R.J., Carrero, J.C., & Lacleite, J.P. Transfection of Platyhelminthes. BioMed Research International. ID 20616 (2015).
17. Ehrhardt, A. *et al.* Episomal vectors for gene therapy. *Current Gene Therapy* **8**, 147–161 (2008).
18. González-Estévez, C., Momose, T., Gehring, W. J. & Saló, E. Transgenic planarian lines obtained by electroporation using transposon-derived vectors and an eye-specific GFP marker. *PNAS* **100**, 14046–14051 (2003).
19. Morales, M. E. *et al.* piggyBac transposon mediated transgenesis of the human blood fluke, *Schistosoma mansoni*. *The FASEB Journ.* **21**, 3479–3489 (2007).
20. Ivics, Z. & Izsvák, Z. The expanding universe of transposon technologies for gene and cell engineering. *Mobile DNA* **1**, 25 (2010).
21. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).
22. Chan, S. Y. & Snow, J. W. Uptake and impact of natural diet-derived small RNA in invertebrates: Implications for ecology and agriculture. *RNA Biology* **14**, 402–414 (2017).
23. Liu, M. A. DNA vaccines: a review. *Journal of Internal Medicine* **253**, 402–410 (2003).
24. Tregoning, J. S., & Kinnear, E. Using plasmids as DNA vaccines for infectious diseases. *Microbiol Spectrum* **2**, <https://doi.org/10.1128/microbiolspec.PLAS-0028-2014> (2014).
25. Palka-Santini, M. *et al.* The gastrointestinal tract as the portal of entry for foreign macromolecules: fate of DNA and proteins. *Mol Gen. Genomics* **270**, 201–215 (2003).
26. Newmark, P. A., Reddien, P. W., Cebria, F. & Sánchez Alvarado, A. Ingestion of bacterially expressed double-stranded RNA inhibits gene expression in planarians. *PNAS* **100**, 11861–11865 (2003).
27. Shih, J. D., Fitzgerald, M. C., Sutherland, M. & Hunter, C. P. The SID-1 double stranded RNA transporter is not selective for dsRNA length. *RNA* **15**, 384–90 (2009).
28. Saleh, M. *et al.* The endocytic pathway mediates cell entry of dsRNA to induce RNAi silencing. *Nat. Cell. Biol.* **8**, 793–802 (2006).
29. Wolff, J. A., Ludtke, J. J., Acsadi, G., Williams, P. & Jani, A. Long-term persistence of plasmid DNA and foreign gene expression in mouse muscle. *Human Molecular Genetics* **1**, 363–369 (1992).
30. McFall-Ngai, M. Divining the essence of symbiosis: insights from the squid-vibrio model. *PLoS Biol.* **12**, e1001783 (2014).
31. Cresci, G. A. & Bawden, E. Gut Microbiome: What we do and don't know. *Nutr. Clin. Pract.* **30**, 734–46 (2015).
32. Gilbert, S. F., Bosch, T. C. & Ledón-Rettig, C. Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat Rev Genet.* **16**, 611–22 (2015).
33. Horn, C. & Wimmer, E. A. A versatile vector set for animal transgenesis. *Dev Genes Evol.* **210**, 630–637 (2000).
34. Handler, A. M. & Harrel, R. A. Germline transformation of *Drosophila melanogaster* with the piggyBac transposon vector. *Insect Molec. Biol.* **8**, 449–457 (1999).
35. Oliveira, L. F. V., Wallau, G. L. & Loreto, E. L. S. Isolation of high quality DNA: a protocol combining rennet and glass milk. *Electronic Journal of Biotechnology* **12**, 1–6 (2009).
36. Krueger, F. Trim Galore!: A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ file, http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ (2015).
37. Hahn, C., Bachmann, L. & Chevreux, B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads - A baiting and iterative mapping approach. *Nucleic Acids Res.* **41**, E129 (2013).
38. Christoff, A. P., Sereia, A. F. R., Boberg, D. R., Moraes, R. L. V., & Oliveira, L. F. V. Bacterial identification through accurate library preparation and high-throughput sequencing. *Neoprosecta Microbiome Technologies*, <http://neoprosecta.com> (2017).
39. Christoff, A. P., Sereia, A. F. R., Hernandez, C. & Oliveira, L. F. V. Uncovering the hidden microbiota in hospital and built environments: New approaches and solutions. *Experimental Biology and Medicine* **244**, 534–542 (2019).
40. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* **29**, 644–52 (2011).
41. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses. *Nucleic Acids Research* **44**, W3–W10 (2016).
42. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature methods* **9**, 357 (2012).
43. Milne, I. *et al.* Using Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics* **14**, 193–202 (2013).

Acknowledgements

We thank Dr. Daniel Graichen, Dr. Henrique B. Ferreira, Dr. Karen Haag and Dr. Lizandra Robe for valuable suggestions and Dr. Charley C. Staats for provide us the pjaf15_H3p-Kat-H3T plasmid. Bioinformatic analysis was performed using the computational facilities CENAPAD/Unicamp (<https://www.cenapad.unicamp.br/>). This project was financially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS/Pronex 16/2551 – 0000499-4).

Author contributions

E.L. and M.R. designed the project, analyzed the data and wrote the manuscript. M.R. prepared the original artwork. All authors have made intellectual contributions to the research project and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-55659-w>.

Correspondence and requests for materials should be addressed to E.L.S.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019

Capítulo 5

***Stenostomum leucops* (Catenulida, Platyhelminthes) has a flexible microbiome in time and space.**

Submitted to the journal *Hydrobiologia*

***Stenostomum leucops* (Catenulida, Platyhelminthes) has a flexible microbiome in time and space**

Rosa, Marcos T^a and Loreto, Elgion LS^{b*}

^a Programa de Pós-Graduação em Biodiversidade Animal, Universidade Federal de Santa Maria, Av. Roraima, 1000; CEP 97105-900; Santa Maria, RS, Brazil. Email: marcostrindadebio@gmail.com;

^b Department of Biochemistry and Molecular Biology, Universidade Federal de Santa Maria, Av. Roraima, 1000; CEP 97105-900; Santa Maria, RS, Brazil. Phone/fax +55 55 32208912 ; Email: elgion@base.ufsm.br

* Corresponding author.

Abstract

Multicellular organisms and their microbiomes can have a restricted and lasting relationship that, sometimes, can be mirrored in their phylogenies, which is called phyllosymbiosis. However, in some organisms, such as freshwater zooplankton, these relationships appear to be more flexible and more easily influenced by the environment. Here, we analysed the microbiome of a freshwater flatworm, *Stenostomum leucops*, sequencing the 16S RNA gene in a metagenomic approach of a strain maintained in the laboratory for 12 years. This strain passed through four different culturing conditions, and the microbiome was characterised for each one. In addition, the microbiome of a wild *S. leucops* population was analysed. The microbiomes were much variable among populations (Shannon index range from 0.43 to 3.05). It was not possible to determine a core microbiome, although *Bacillus mycoides*, *Pseudomonas*, *Flectobacillus*, and *Ralstonia pickettii* were the predominant bacteria in populations under stable conditions. In culturing conditions, in which the water was contaminated by iron, *Rhodospirillum rubrum*, which is a bacteria involved in iron reduction, was the prevalent species. Our findings were concordant with other studies on freshwater zooplankton. The microbiomes were very flexible and influenced by the environment.

Keywords: phyllosymbiosis, microbiome, iron contamination, bacterial succession.

Introduction

Multicellular eukaryotes always live in an environment replete of microbes and interact with them. Often, these interactions are strong, and the involved organisms are completely dependent on one another. In extreme cases, for example, the bacteria become intracellular components of the host, as in several annelids, molluscs, and platyhelminthes, where bacteria are maintained in a specialised cell organelle, the trophosome, supplying energy for the host (Dubilier et al., 2008). In other cases, the microbe–host association involves the exchange of different materials, such as

nutrients, probiotics, and secondary metabolites (Phillips et al., 2012; Moeller et al., 2014). In fact, these associations, at least for some organisms, are relevant for their fitness and evolution, once they are maintained along evolutionary time and can be detected by phylosymbiosis, as defined by Lin & Bordenstein (2020) as “microbial community relationships that recapitulate the phylogeny of their host.”

Phylosymbiosis is observed in several animal taxa. For example, in scleractinian corals (Pollock et al., 2018; Gault et al., 2021), sponges (Easson & Thacker, 2014), molluscs (Huot et al., 2019), and different insects, such as bees (Figuerola et al., 2021), ants (Sanders et al., 2014), wasps, and *Drosophila*, (Brooks et al., 2016). Additionally, it has also been described in mammals, such as mice, bats, and primates (Phillips et al., 2012; Moeller et al., 2014; Sanders et al., 2014). However, for several organisms, a phylosymbiosis signal was not found, such as in birds (Hird et al., 2015; Kropáčková et al., 2017), Rotifera, and Crustacea (Eckert et al., 2020, 2021). This last study showed that freshwater zooplankton had no signal of phylosymbiosis; in contrast, the microbiome of these organisms is highly flexible and can be influenced by the external environment. Therefore, the degree of functional and evolutionary constraints involving the microbiome and its hosts' relationships is yet an open question, and in this moment, more research covering the most variable kinds of organisms are necessary to answer this question.

Catenulida is a basal class of Platyhelminthes and is represented by around 100 described species of small freshwater flatworms ranging in length from 0.5 to 5 mm (Noreña et al., 2005). They reproduce asexually by a process called paratomy, in which the development of structures typical of the anterior region in posterior parts of the body occurs, and after complete development, fission produces two or more individuals (Rieger, 1986; Noreña et al., 2005). Among Catenulida, the *Paracatenula* have a restricted relationship with some bacteria. Chemoautotrophic bacteria of the genus *Candidatus* are an obligatory symbiont in the trophosomes of *Paracatenula*, corresponding to half of the biomass of these worms and being the primary energy source (Oliver et al., 2018). Another Catenulida, *Stenostimum leucops*, also has a special interaction with its microbiome, which is suggested to modulate phenotypic plasticity (Rosa et al., 2015). Additionally, these worms can express genes maintained by their microbiome, at least the worms express green fluorescent protein (GFP) when hosting bacteria having plasmids with the *GFP* gene (Rosa & Loreto, 2019). However, none of the studies have focused on the microbiome of Catenulida representatives to determine how constant or flexible the bacterial community is in these worms.

In this study, we addressed the microbiome associated with *S. leucops* using a descriptive approach, in which the microbiome was assessed in time, looking at a laboratory strain during a

period of six years that experienced four different culturing processes. Additionally, the microbiome of the laboratory population was compared with one population recently collected in the wild.

Materials and Methods

Animals used and their maintenance

Two different sources of *S. leucops* were used in this work: i) a wild population freshly collected in a pond at the Federal University of Santa Maria, Santa Maria, Brazil (53°17'W; 29°28'S) and ii) a clonal isolate maintained in the laboratory for more than 12 years, at 28°C, in reconstituted water and fed with powdered milk. The morphological, DNA barcode, and reproductive characteristics of these strains, as well as maintenance procedures, were described in Rosa et al. (2015) and Rosa & Loreto (2021). The microbiome of the laboratory strain was established over time in four different moments (Fig 1.5).

During maintenance in the laboratory, some problems occurred. In the first 8 years, the worms were maintained with reconstituted water, made with distilled water, and the animals were easily maintained (Fig 1.5). Selection was performed to obtain cultures only with *S. leucops* and its microbiome. In 2017, the strain suddenly passed through a bottleneck and became more difficult to maintain. To recover strain vigour, we supplemented the culture with water from a dam, which was centrifuged at 1000 rpm for 5 minutes, and the supernatant was added to the culture, aiming to improve the bacterial species available for the worms' microbiota. The purpose of centrifugation was to prevent other worms from contaminating the culture. Powdered fish food was also included in the worms' diet. Iron contamination was detected as a cause of the bottleneck, and in 2018, the culture started to be maintained using deionised water to make the reconstituted water. The vigour of the strain increased. In 2020, due to the Sars-COV2 pandemic, we decided to maintain a subculture at home using tap water.

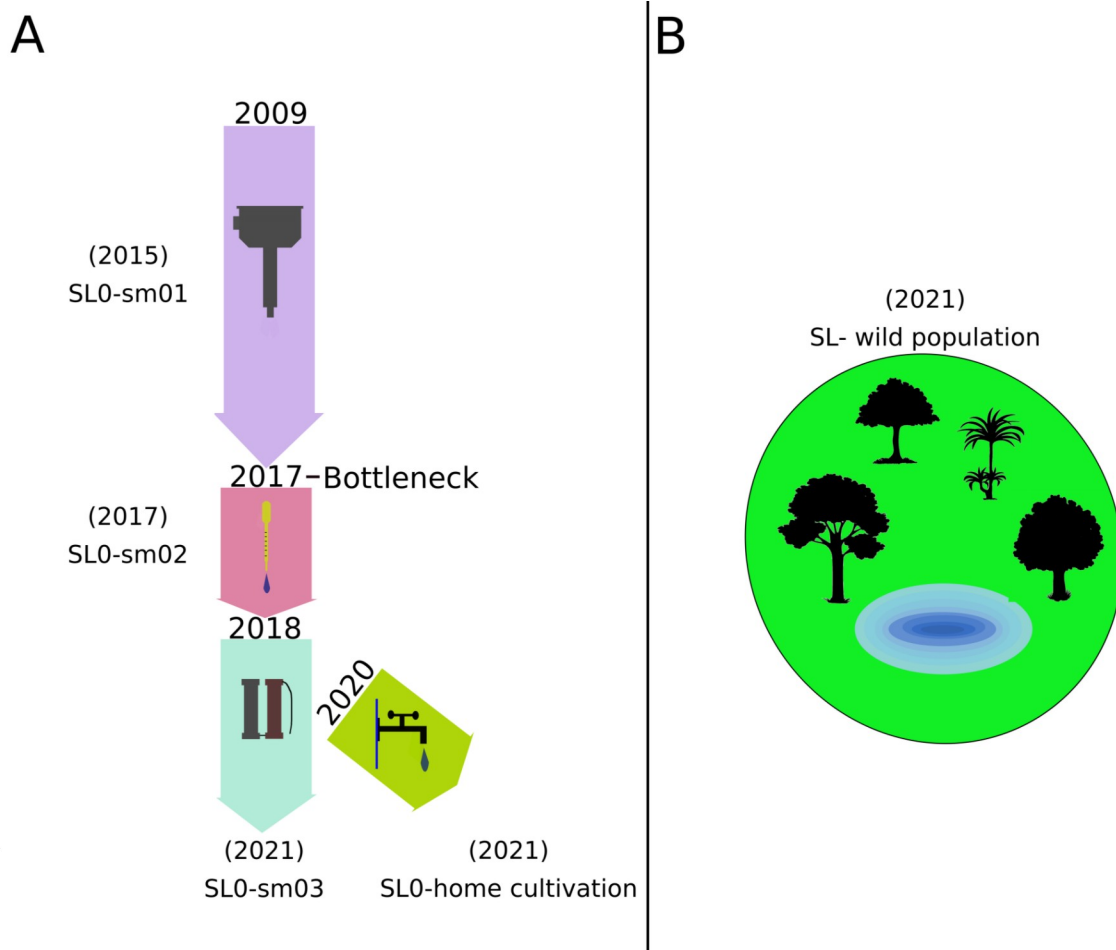


Figure 1.5: A) A laboratory *Stenostomum leucops* isolate and its maintenance phases: 2009–2017, maintained with distilled water under selection for a cleaner culture (SLO-sm01); in 2017, occurred a bottleneck occurred, and “new” bacteria sources were added to rescue culture vigour (SLO-sm02); in 2018, iron contamination in the water was observed and the culture became to be maintained with deionised water (SLO-sm03). In 2020, a subpopulation was maintained at home using tap water (SLO-hc). The microbiomes were established in 2015 (SLO-sm01), 2017 (SLO-sm02), and 2021 (SLO-sm03 and SLO-home cultivation). B) A wild population was collected in a pond at the Federal University of Santa Maria, Santa Maria, Brazil (53°17’W; 29°28’S) in September 2009. These specimens were identified following Noreña *et al.* (2005) and by DNA barcoding following Rosa *et al.* (2005).

Microbiome analyses

For microbiome analyses, total DNA was isolated from a pool of 100 individuals using a protocol described by Oliveira *et al.* (2009). The worms were collected in the culture plates using a micropipette and first transferred to a drop of reconstituted water to minimise contamination for environmental microorganisms and finally transferred to a microtube containing extraction buffer for DNA isolation. Bacteria identification was performed using high-performance sequencing of the V3/V4 regions of the 16S rRNA gene. The preparation of the libraries followed a proprietary protocol (Neopropecta Microbiome Technologies, Brazil). Amplification was performed with

primers for the V3–V4 region of the 16S rRNA gene, 341F (CCTACGGGRSGCAGCAG; Wang & Qian, 2009) and 806R (GGACTACHVGGGTWTCTAAT; Caporaso et al., 2012). The libraries were sequenced using MiSeq Sequencing System equipment (Illumina Inc., USA). A MiSeq Reagent Kit v3 (600 cycles) was used for sequencing to produce paired end reads.

The microbiome sequencing reads are available under the following NCBI accession numbers: SAMN12561444 (*Stenostomum leucops* SL0-sm01); SAMN12561445 (*S. leucops* SL0-sm02); SAMN24102714 (*S. leucops* SL0-sm03); SAMN24102715 (*S. leucops* SL0-hc); and SAMN24102716 (*S. leucops* SL0-wp).

For bioinformatics analysis, first the reads were processed in Fastqc (Andrews 2010), and only reads with a quality higher than 30 were utilised. The first 5 nucleotides and the last 10 to 25 were removed using TrimGalore 0.6.7 (<https://github.com/FelixKrueger/TrimGalore/releases/tag/0.6.7>).

For microbiome classification using 16S sequences, the reads were compared using the Blast+ tool (Camacho et al., 2009) with the 16S RefSeq NCBI databank (<https://ncbiinsights.ncbi.nlm.nih.gov/2020/02/21/rrna-databases/>). The conditions used for local blastn were a minimum of 75% identity, utilising the 5 best hits for each sequence. The tables were ordered to remove more than one classification for each sequence; selection was made by the best score and lowest e-value. The predominant microbiome in each population was considered the species that showed more than 100 reads in the NSG sequence.

Determination of iron content in water

For iron determination, an internal calibration curve with a linear calibration range from 1 to 100 µg/L was set up using the SCP Science Kit. The measuring instrument was an inductively coupled plasma optical emission spectrometer (model Optima 4300 DV, Perkin Elmer, USA). The monitored wavelengths were 238.204 nm and 239.562 nm.

Results

Microbiome analyses

The number and size of reads used for microbiome estimation for each population can be found in Supplementary Table 1. The smaller dataset for SLO-sm01 had 16905 reads, and the bigger dataset for SLwild_population contained 87698 reads. The analysis of microbiome diversity (Table 1.5) showed that SLO_sm01 had a reduced diversity (Shannon index = 0.43), and the other

populations had a higher diversity (1.36 to 3.05). The populations showing higher diversity had distinct bacterial communities (β diversity index = 3.84). However, the evenness index (J') ranging from 0.22 to 0.47 showed that there were a few highly dominant species and a large number of rare species.

	SLO_sm01	SLO_sm02	SLO_sm03	SLO_hc	SLO_wp
H'	0.43	3.05	1.36	2.16	2.13
J'	0.22	0.47	0.24	0.32	0.37
β	3,839907193				

Table 1.5: Shannon diversity index (H'); evenness index (J') and β diversity index for *Stenostomum leucops* populations analyzed.

The microbiomes found in *Stenostomum leucops* in the laboratory populations followed up over a period of 12 years or compared with a recently collected wild population were greatly variable in number and content of bacterial species. A full list of species, their taxonomy, and the number of reads obtained in the sequencing for each population can be found in the Supplementary spreadsheet. These results can be visualised in Figures 2.5 and 3.5.

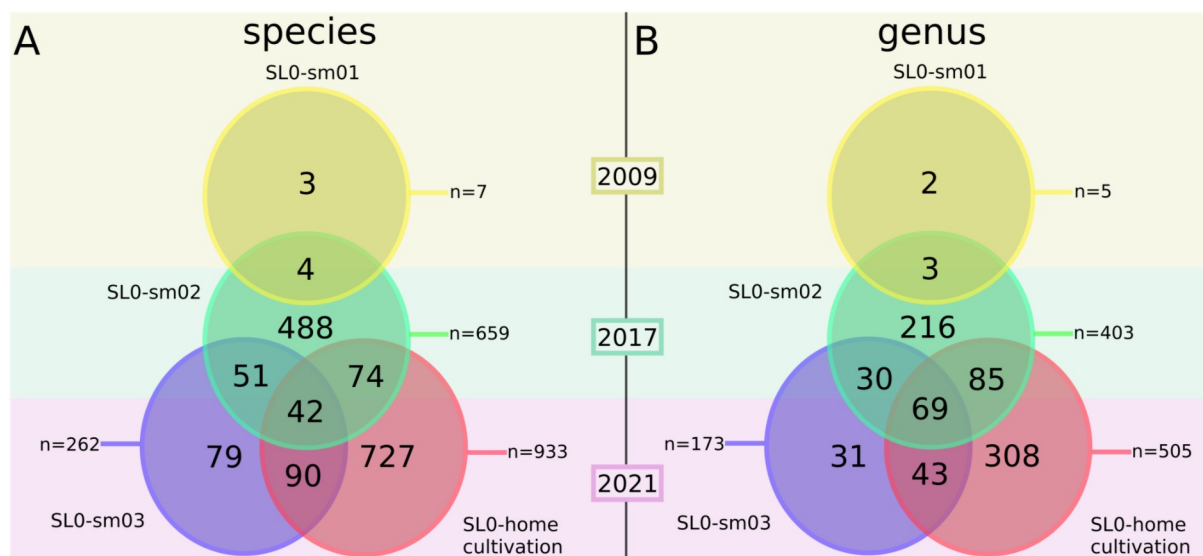


Figure 2.5: Venn diagram of the total bacterial species (A) and genus (B), estimated by 16S rRNA gene sequencing, found in a *S. leucops* laboratory strain over time and under differential cultivation conditions (home/laboratory). The circle colour indicates = yellow—SLO-sm01; green—SLO-sm02; purple—SLO-sm03; and pink—SLO-hc.

The laboratory population (SLO_sm01), when analysed after 6 years of maintenance, had seven species belonging to 5 genera (Fig 2.5). After a bottleneck, we were forced to add new bacteria so as not to lose the worm's strain (SLO_sm02); the number of bacterial species increased to 659, four of which were shared with the strain before bottleneck. After four years of laboratory

maintenance (SLO_sm03), this number decreased to 262, of which 93 were shared with SLO_sm02. Maintenance for two years using tap water (home cultivation) greatly improved the microbiome, increasing to 933 bacterial species. The number of bacteria shared with SLO_sm02 was 116 and 132 with SLO_sm03. The comparison of microbiomes of extant populations, the laboratory population (SLO_sm03), the home cultivation (SLO_hc) population, and the wild population (SLO_wp), can be found in Fig 3.5. A total of 63 species were shared in three populations. The wild population had 294 species, and 125 were shared with SLO_hc and 93 with SLO_sm03.

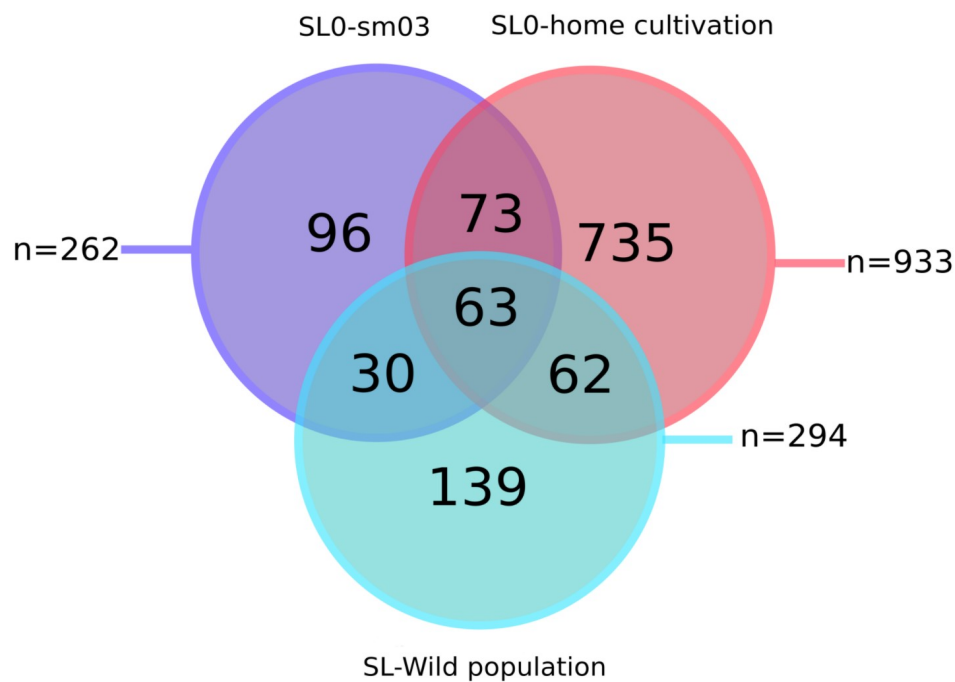


Figure 3.5: Venn diagram of the total bacterial species, estimated by 16S rRNA gene sequencing, found in the *S. leucops* laboratory strain (purple), the population cultivated with tap water SLO-hc (home cultivation-pink), and a wild population (SLO-wp—blue).

To characterise the predominant microbiome in populations, only the bacterial species that showed more than 100 reads in the sequencing analysis were analysed (Table 2.5). In general, these species corresponded to 96 to 98% of obtained reads. For the SLO_sm02 population, the predominant microbiome corresponded to 82% of reads. The predominant bacteria belong, in general, to the Proteobacteria phylum, followed by Firmicutes. None of the predominant species were shared by all populations. However, *Bacillus mycoides* was the predominant bacteria in the first laboratory population (SLO_sm01) and was absent among the predominant species during the bottleneck (SLO_sm02) but returned to being the most predominant in the current population (SLO_sm03). In addition, this bacterial species was most prevalent in the wild population

(SLO_wp). *Bacillus mycoides* was also present in the home cultivation population (SLO_hc), although it corresponds to only 0.7% of reads.

SLO-sm01			SLO-sm02			SLO-sm03			SLO-home cultivation			SLO-wild population		
%	Species	Phylum	%	Species	Phylum	%	Species	Phylum	%	Species	Phylum	%	Species	Phylum
83.5	<i>Bacillus mycoides</i>	Firmicutes	42.7	<i>Rhodferax ferrireducens</i>	Proteobacteria	66.2	<i>Bacillus mycoides</i>	Firmicutes	29.7	<i>Pseudomonas libanensis</i>	Proteobacteria	33.8	<i>Bacillus mycoides</i>	Firmicutes
14.9	<i>Escherichia fergusonii</i>	Proteobacteria	10.1	<i>Rhizobacter profundus</i>	Proteobacteria	10.2	<i>Kinnetrieta asaccharophila</i>	Proteobacteria	24.4	<i>Flectobacillus roseus</i>	Bacteroidetes	22.6	<i>Aeromonas dhakensis</i>	Proteobacteria
			6.1	<i>Granulicella acidiphila</i>	Acidobacteria	9.9	<i>Ralstonia pickettii</i>	Proteobacteria	15.4	<i>Pseudomonas alcaligenes</i>	Proteobacteria	9.9	<i>Pseudomonas plecoglossicida</i>	Proteobacteria
			3.5	<i>Lysoabacter capsici</i>	Proteobacteria	4.9	<i>Pseudomonas libanensis</i>	Proteobacteria	13.6	<i>Ralstonia pickettii</i>	Proteobacteria	8.4	<i>Comamonas testosteroni</i>	Proteobacteria
			3.1	<i>Curvibacter delicatus</i>	Proteobacteria	1.9	<i>Bacillus flexus</i>	Firmicutes	5.8	<i>Flectobacillus rhizosphaerae</i>	Bacteroidetes	7.1	<i>Pseudomonas montelii</i>	Proteobacteria
			1.5	<i>Ensifer adhaerens</i>	Proteobacteria	1.2	<i>Bacillus aryabhatai</i>	Firmicutes	1.1	<i>Pseudomonas gessardii</i>	Proteobacteria	5.3	<i>Aeromonas caviae</i>	Proteobacteria
			1.2	<i>Halodesulfobivrio spirochaetisodalis</i>	Proteobacteria	0.9	<i>Ideonella sakaensis</i>	Proteobacteria	0.8	<i>Pseudomonas stutzeri</i>	Proteobacteria	4.0	<i>Pseudomonas putida</i>	Proteobacteria
			1.1	<i>Thermonaerobaculum aquaticum</i>	Acidobacteria	0.6	<i>Pseudomonas korensis</i>	Proteobacteria	0.7	<i>Bacillus mycoides</i>	Firmicutes	1.2	<i>Pseudomonas benzenivorans</i>	Proteobacteria
			1.1	<i>Tibeticola sediminis</i>	Proteobacteria	0.6	<i>Acidovorax soli</i>	Proteobacteria	0.7	<i>Pseudomonas parafulva</i>	Proteobacteria	1.0	<i>Pseudomonas flavescens</i>	Proteobacteria
			1.0	<i>Hyphomicrobium vulgare</i>	Proteobacteria	0.5	<i>Pseudomonas guetzneii</i>	Proteobacteria	0.6	<i>Pseudomonas compositi</i>	Proteobacteria	0.8	<i>Ralstonia pickettii</i>	Proteobacteria
			1.0	<i>Bdellovibrio ovovorans</i>	Proteobacteria	0.5	<i>Pseudomonas parafulva</i>	Proteobacteria	0.5	<i>Pseudomonas knackmussii</i>	Proteobacteria	0.8	<i>Pseudomonas segetis</i>	Proteobacteria
			0.8	<i>Nocardioides salisibiostraticola</i>	Actinobacteria	0.4	<i>Pseudomonas alcaligenes</i>	Proteobacteria	0.5	<i>Pseudomonas korensis</i>	Proteobacteria	0.7	<i>Pseudomonas vranovensis</i>	Proteobacteria
			0.8	<i>Hyphomicrobium zavarzinii</i>	Proteobacteria	0.3	<i>Tibeticola sediminis</i>	Proteobacteria	0.5	<i>Acidovorax soli</i>	Proteobacteria	0.6	<i>Flectobacillus roseus</i>	Bacteroidetes
			0.7	<i>Nitrospira</i>	Nitrospirae	0.3	<i>Aquabacterium parvum</i>	Proteobacteria	0.5	<i>Pseudomonas migulae</i>	Proteobacteria	0.5	<i>Pseudomonas kilonensis</i>	Proteobacteria
			0.7	<i>Panacagninomonas persipica</i>	Proteobacteria	0.2	<i>Delftia acidovorans</i>	Proteobacteria	0.4	<i>Pseudomonas aeruginosa</i>	Proteobacteria	0.3	<i>Flectobacillus rhizosphaerae</i>	Bacteroidetes
			0.7	<i>Nevskia lacus</i>	Proteobacteria	0.1	<i>Rhodoflex ferrireducens</i>	Proteobacteria	0.3	<i>Pandorea verucati</i>	Proteobacteria	0.3	<i>Pseudomonas nitroreducens</i>	Proteobacteria
			0.7	<i>Brevundimonas alba</i>	Proteobacteria				0.2	<i>Pseudomonas fluvialis</i>	Proteobacteria	0.2	<i>Pseudomonas libanensis</i>	Proteobacteria
			0.7	<i>Aquaspirillum serpens</i>	Proteobacteria				0.2	<i>Cupriavidus necator</i>	Proteobacteria	0.2	<i>Pseudomonas balearica</i>	Proteobacteria
			0.6	<i>Tabrizicola aquatica</i>	Proteobacteria				0.2	<i>Ralstonia insidiosa</i>	Proteobacteria	0.2	<i>Pseudomonas mangrovi</i>	Proteobacteria
			0.6	<i>Lacisediminimonas profundus</i>	Proteobacteria				0.2	<i>Pseudomonas segetis</i>	Proteobacteria	0.2	<i>Pseudomonas lundensis</i>	Proteobacteria
			0.5	<i>Lacobacterium aquatile</i>	Proteobacteria				0.1	<i>Fusobacterium mortiferum</i>	Fusobacteria	0.1	<i>Pseudomonas parafulva</i>	Proteobacteria
			0.5	<i>Rhizobacter fulvus</i>	Proteobacteria									
			0.5	<i>Rhizobacter subsaxonicus</i>	Proteobacteria									
			0.4	<i>Desulfobivrio vietnamensis</i>	Proteobacteria									
			0.4	<i>Lentimicrobium saccharophilum</i>	Bacteroidetes									
			0.4	<i>Inquilinus limosus</i>	Proteobacteria									
			0.3	<i>Chryseolinea soli</i>	Bacteroidetes									
			0.3	<i>Ideonella paludis</i>	Proteobacteria									
98.4	% of reads		82.0	% of reads		98.6	% of reads		96.3	% of reads		98.1	% of reads	

Table

2.5 – Species and respective phyla of predominant microbiome found in different *Stenostomum leucops* populations% = correspond to percentage of sequenced reads; color mark: these species are shared by SLO-sm01; SLO-sm03, SLO-home, SLO-wild; SLO-sm03, SLO-home, SLO-wild; SLO-sm02 and SLO-sm03; SLO-sm03, SLO-home; SLO-home, SLO-wild.

In the home cultivation population (SLO_hc), the most predominant species were *Pseudomonas libanensis*, *Flectobacillus roseus*, *Pseudomonas alcaligenes*, *Ralstonia pickettii*, and *Flectobacillus rhizosphaera*, and these species are also among the predominant species of other extant populations (SLO_sm03 and SLO_wp). The overrepresented bacterial species during the bottleneck was *Rhodoflex ferrireducens* (42.7% of reads). Although this species is present in the current population (SLO_sm03), it corresponds to only 0.1% of reads (Table 2.5).

Determination of iron content in water

During laboratory maintenance, we experienced difficulties in growing the worms, and we associated this fact with the iron contamination in the pipeline water. Furthermore, during this period, the most abundant bacteria in the worm microbiome was *Rhodoflex ferrireducens*, which is involved in iron reduction; thus, we determined the iron content in some water sources (Table 3.5).

In the house tap water and laboratory deionised water, the amount of iron was less than 5 $\mu\text{g L}^{-1}$; the laboratory pipeline water and dam water contained more than 800 $\mu\text{g L}^{-1}$.

<i>Sample</i>	<i>Iron concentration - $\mu\text{g L}^{-1}$</i>
Laboratory (pipeline water)	813
Laboratory deionized water	< 5
House (tap water)	< 5
Dam water (where was collected the wild population)	987

Table 3.5: Determination of Iron Content in Water.

Discussion

Today, it is clear that the microbiome plays a relevant role in the biology of its host (Bordenstein & Theis, 2015), although our current knowledge suggests that different organisms respond differently to this host–microbiome association regarding physiology, stress tolerance, fitness, and evolvability. Organisms, such as corals, have a strong association with their microbiome but also show some flexibility in this association as a way to quickly adapt to environmental changes (Voolstra & Ziegler 2020). Studies have focused on freshwater zooplankton, although few studies have studied Crustacea (Cladocera and Copepoda) and Rotifera, which have highly flexible microbiomes (Eckert et al., 2020; Eckert et al., 2021; Akbar et al., 2022). The freshwater flatworm, *S. leucops*, showed a flexible microbiome, similar to the other freshwater zooplankton studied thus far.

The microbiome of the *Stenostomum leucops* laboratory strain followed in time was very flexible, in comparison with a wild population. In general, most bacteria were present in low frequency and were likely obtained from the environment. However, when analysing the most frequent bacteria, a core microbiome was not found, although for the “stable” laboratory populations, which were easy to maintain and proliferative (SLO_sm01 and SLO_sm03) and for the wild population (SLO_wp), the most abundant bacterial species was *Bacillus mycoides*. However, in the population maintained at home (SLO_hc), which was also very proliferative, the most abundant bacteria were *Pseudomonas* and *Flectobacillus*. *Bacillus mycoides*, although present, had a low abundance. For the current laboratory population (SLO_sm03), home cultivation population, and wild population, bacteria *Ralstonia pickettii* was also abundant.

The population SLO_sm02, which went through a severe bottleneck due to iron contamination in the pipeline water in the laboratory building, the predominant bacteria was *Rhodoferrax*

ferrireducens. This bacterial species can support growth using Fe (III) reduction (Finneran et al., 2003). However, during the period in which this bacterial species was predominant, the *Stenostomum* population was difficult to maintain; although having a richer microbiome biodiversity, the bacteria abundant in the “stable” laboratory populations were absent in SLO_sm02. Alterations of microbiota by different chemicals can negatively influence host fitness and physiology (Brüssow, 2015; Jin et al., 2017; Pennycook & Scanlan, 2021). The water in the pond in which the wild population (SLO_wp) was collected also showed a high iron concentration, which was three times higher than the limit for dissolved iron in drinking water (World Health Organization, 2008). However, *Rhodospirillum rubrum* was not present in the SLO_wp sample. Thus, these results suggest that iron alone was not the stressor of *Stenostomum*, but the microbiome affected by this metal promotes difficulties in laboratory maintenance of this species.

In conclusion, the *S. leucops* microbiome was variable over time in both a laboratory strain and a wild population. However, we can register here as a case report that the wild population and the laboratory population, in moments that were “easy” to maintain, shared several bacteria, such as *Bacillus mycoides*, *Pseudomonas*, *Flectobacillus*, and *Ralstonia pickettii*. These last results indicate that *S. leucops* is a promissory organism to perform experiments with microbiome shift.

Acknowledgements

We thank Neoprospecta Microbiome Technologies for assistance in the sequencing. This study was supported by research grants and fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico— CNPq (421255/2018-8 and 305613/2020-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

References:

- Akbar S, Gu L, Sun Y, Zhang L, Lyu K, Huang Y & Yang Z. 2022. Understanding host-microbiome-environment interactions: Insights from *Daphnia* as a model organism. *Science of The Total Environment* 808: 152093
- Andrews S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- Bordenstein SR & Theis KR. 2015. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biology* 13: e1002226.
- Brooks AW, Kohl KD, Brucker RM, van Opstal EJ & Bordenstein SR. 2016. Phylosymbiosis: Relationships and Functional Effects of Microbial Communities across Host Evolutionary History. *PLoS Biology* 14: e2000225.
- Brüssow H. 2015. Growth promotion and gut microbiota: insights from antibiotic use. *Environmental Microbiology* 17:2216-2227.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K & Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.
- Caporaso J, Lauber C, Walters, et al. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *The ISME Journal* 6:1621–1624
- Dubilier N, Bergin C & Lott C. 2008. Symbiotic diversity in marine animals: The art of harnessing chemosynthesis. *Nature Review Microbiology* 6:725–740.
- Easson CG & Thacker RW. 2014. Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Frontiers in Microbiology* 5:532.
- Eckert E, Amalfitano S, Di Cesare A, Manzari C, Corno G & Fontaneto D. 2020. Different substrates within a lake harbour connected but specialised microbial communities. *Hydrobiologia* 847: 1689-1704.

Eckert EM, Anicic N & Fontaneto D. 2021. Freshwater zooplankton microbiome composition is highly flexible and strongly influenced by the environment. *Molecular Ecology* 30:1545–1558.

Figuerola LL, Maccaro JJ, Krichilsky E, Yanega D & McFrederickc QS. 2021. Why Did the Bee Eat the Chicken? Symbiont Gain, Loss, and Retention in the Vulture Bee Microbiome. *MBio* 12: e02317-21.

Finneran KT, Johnsen CV & Lovley DR. 2003. *Rhodoferax ferrireducens* sp. nov., a psychrotolerant, facultatively anaerobic bacterium that oxidizes acetate with the reduction of Fe(III). *International Journal of Systematic and Evolutionary Microbiology*. 53:669-673.

Gault, JA, Bentlage B, Huang D & Kerr AM. 2021. Lineage-specific variation in the evolutionary stability of coral photosymbiosis. *Science Advances* 7: eabh4243.

Hird SM, Sánchez C, Carstens BC & Brumfield RT. 2015 Comparative gut microbiota of 59 neotropical bird species. *Frontiers in Microbiology* 6: 1403.

Huot C, Clerissi C, Gourbal B, Galinier R, Duval D & Toulza E. 2020. Schistosomiasis Vector Snails and Their Microbiota Display a Phylosymbiosis Pattern. *Frontiers in Microbiology* 10: 3092.

Jin Y, Wu S, Zeng Z & Fu Z. 2017. Effects of environmental pollutants on gut microbiota. *Environmental Pollution* 222: 1-9.

Kropáčková L, Těšický M, Albrecht T, Kubovčíak J, Čížková D, Tomášek O, Martin JF, Bobek L, Králová T, Procházka P & Kreisinger J. 2017. Codiversification of gastrointestinal microbiota and phylogeny in passerines is not explained by ecological divergence. *Molecular Ecology* 26:5292-5304.

Lim SJ & Bordenstein SR. 2020. An introduction to phylosymbiosis. *Proceeding of Royal Society. B* 287: 20192900.

Moeller AH, Li Y, Ngole EM, Ahuka-Mundeké S, Lonsdorf EV, Pusey AE, Peeters M, Hahn BH & Ochman H. 2014. Rapid changes in the gut microbiome during human evolution. *Proceeding of National Academy of Science. USA* 111: 431–16 435.

Noreña C, Damborenea C & Brusa F. 2005. A taxonomic revision of South American species of the genus *Stenostomum* O. Schmidt (Platyhelminthes: Catenulida) based on morphological characters. *Zoological Journal of the Linnean Society* 144: 37-58.

Oliveira LFV, Wallau GL & Loreto ELS. 2009. Isolation of high quality DNA: a protocol combining rennet and glass milk. *Electronic Journal of Biotechnology* 12, 1–6.

Oliver J, Seah BKB, Tietjen M, Leisch N, Liebeke M, Kleiner M, Berg JS & Gruber-Vodicka HR. 2018. Chemosynthetic symbiont with a drastically reduced genome serves as primary energy storage in the marine flatworm *Paracatenula*. *Proceeding of National Academy of Science. USA* . 116: 8505-8514.

Pennycook JH & Scanlan PD. Ecological and Evolutionary responses to Antibiotic Treatment in the Human Gut Microbiota. 2021. *FEMS Microbiology Review* 45:fuab018.

Phillips CD, Phelan G, Dowd SE, McDonough MM, Ferguson AW, Delton Hanson J, Siles L, Ordóñez-Garza N, San Francisco M & Baker RJ. 2012. Microbiome analysis among bats describes influences of host phylogeny, life history, physiology and geography. *Molecular Ecology* 21:2617–2627.

Pollock FJ, McMinds R, Smith S, Bourne DG, Willis BL, Medina M, Thurber RV & Zaneveld JR. 2018. Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nature Communications* 9: 4921.

Rieger RM. 1986. Asexual reproduction and the turbellarian archetype. *Hydrobiologia*. 132:35–45.

Rosa MT & Loreto ELS. 2019. The Catenulida flatworm can express genes from its microbiome or from the DNA it ingests. *Scientific Report* 9:19045.

Rosa MT & Loreto ELS. 2021. Revisiting the regeneration of *Stenostomum leucops* (Catenulida, Platyhelminthes), Invertebrate Reproduction & Development, DOI: 10.1080/07924259.2021.1994475

Rosa MT, Pereira CM, Ragagnin GT & Loreto ELS. 2015. *Stenostomum leucops* Duges, 1828 (Platyhelminthes, Catenulida): A putative species complex with phenotypic plasticity. *Pap. Avulsos Zool.* 55:375–383.

Sanders JG, Powell S, Kronauer DJ, Vasconcelos HL, Frederickson ME & Pierce NE. 2014. Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. *Molecular Ecology* 23:1268–1283.

Voolstra CR & Ziegler M. 2020. Adapting with Microbial Help: Microbiome Flexibility Facilitates Rapid Responses to Environmental Change. *Bioessays* 42:e2000004.

Wang Y & Qian P-Y. 2009 Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies. *PLoS ONE* 4: e7401.

World Health Organization. *Iron in drinking-water*. Background document for preparation of WHO Guidelines for drinking-water quality, 2008. Geneva, World Health Organization (WHO/SDE/WSH/03.04/8) 390. Website http://www.who.int/water_sanitation_health/dwq/chemicals/iron.pdf.

Supplementary Material:

*Supplementary table 16S sequencing and legends supplementary table spreadsheets.

(https://www.dropbox.com/s/adwvjjoxxylqq7k/Suppl_material.pdf?dl=0)

*Supplementary table spreadsheets.

(https://www.dropbox.com/scl/fi/9tinb3ucqiv16isjbnpi9/Supp_spreadshets_microbiome.xls?dl=0&rlkey=715mhjf1871nu744phvbc8s1y)

Capítulo 6

Análise do genoma e transcriptoma de *Stenostomum leucops*.

Resultados preliminares e perspectivas

Até o presente momento não há publicações caracterizando o genoma de nenhum *Catenulida*. Neste sentido, buscamos sequenciar e caracterizar o genoma de *S. leucops*. Primeiramente fizemos um sequenciamento Illumina, mas devido ao número de *reads* gerados (baixa cobertura) e a proporção elevada de regiões repetitivas encontrada neste genoma, não obtivemos êxito em uma boa caracterização e anotação do genoma sequenciado. Em uma tentativa de superar esta limitação, resolvemos fazer um sequenciamento de *long reads*, utilizando a plataforma Oxford Nanopore. Além disso, também obtivemos um transcriptoma de vermes em estágios de um e dois zooides, em que objetivamos estudar a expressão gênica diferencial nesses estágios. Para este estudo empregando a plataforma IonTorrent S5.

As análises dos genomas e transcriptomas não puderam ser finalizadas. O principal entrave que encontramos, para que a análise não fosse terminada antes da defesa desta tese, foi a dificuldade de obter amostras de DNA em quantidade e qualidade apropriada para o sequenciamento Nanopore. Um total de 10 microgramas de DNA, com tamanho médio de 40 kb, são solicitados para o sequenciamento. Para obter este material foi necessário aproximadamente 30.000 vermes. Lembramos que a diversidade genética, que poderia ser um problema para montagem de um genoma, partindo de uma amostra empregando um número tão grande de indivíduos, neste caso não é significativo, uma vez que é uma isolinhagem de reprodução assexuada – clonal.

A pandemia Covid-19 limitou o acesso ao laboratório por um longo período, assim, após o primeiro envio de amostras para o sequenciamento, esta foi rejeitada (pequena quantidade e qualidade não apropriada). Tivemos de reiniciar o cultivo de organismos para nova extração, dentro dos limites institucionais para o acesso ao laboratório. Somente a partir do segundo envio, após meses de cultivo e extração de DNA é que conseguimos as sequências desejadas.

Apresentamos a seguir os resultados preliminares das análises do genoma e transcriptoma. As análises necessitam continuidade para que possamos caracterizar e anotar o genoma de *S. leucops* e trazer informações funcionais sobre o processo de expressão gênica durante o desenvolvimento dos zooides.

Materiais e métodos

Sequenciamento e montagem Illumina

O DNA genômico foi extraído utilizando o protocolo de *glass milk* (Oliveira et al. 2009), partindo de uma população com 100 organismos de *S. leucops* SI-SM01. O DNA obtido passou por uma amplificação através do kit Genomiphi™ V2 DNA Amplification (Thermo), para a obtenção de

aproximadamente 2 µg de DNA genômico. O sequenciamento foi realizado no Laboratório de Genômica e Bioinformática do LNCC/MCT, Petrópolis/RJ, em uma plataforma Illumina miseq.

Os *reads* passaram pelo teste de qualidade através do software Fasqc (Andrews 2010), para a remoção de adaptadores e *reads* de qualidade inferior a 30 (recomendado para tecnologia Illumina), este *trimming* foi feito pelo software Trimgalore (Krueger 2015). A montagem do genoma foi implementada por três softwares: SPAdes (Bankevich et al. 2012), AbySS (Nielsen et al. 2009) e SOAPdenovo (Li et al. 2010), através da metodologia múltiplos K-mers. Os *reads* foram submetidos a vários *K-mers* de montagem: 15; 30; 45; 60, para selecionar o *K-mer* que apresente os melhores parâmetros de montagem.

Sequenciamento e montagem do transcriptoma

O RNA total foi extraído com Trizol (Invitrogen ®) e clorofórmio. Utilizou-se cerca de 300 organismos para cada réplica e tratamento. Utilizamos três réplicas e dois tratamentos (estágios). Os organismos foram capturados individualmente, com o uso de uma micropipeta, e visualizados com auxílio de um microscópio invertido com aumento de 40X. Os organismos foram separados em dois grupos com relação a formação de zooides. Tratamento 1, denominado “II zooides”, composto por organismos com formação de fossos ciliares na região mediana do corpo, apresentando dois zooides. Tratamento 2, denominado “I zoóide”, formado por organismos sem alongação ou formação de zoóide. Foram realizadas triplicatas técnicas. Para preparação das bibliotecas, o mRNA foi enriquecido com o kit de purificação Dynabeads® mRNA (Invitrogen), e utilizado o Total Ion RNA-Seq v2 kit (Thermo Fisher Scientific) no sistema *One Touch* (Thermo Fisher Scientific). O sequenciamento foi realizado na plataforma IonTorrent S5 sequencer (Thermo Fisher Scientific) usando o Ion 540 TM Kit-OT2 e Ion 540 TM Chip.

Os *reads* passaram pelo teste de qualidade através do software Fasqc (Andrews 2010), para a remoção de adaptadores e *reads* de qualidade inferior a 20 (recomendado para tecnologia RNA-seq IonTorrent), este *trimming* foi feito pelo software Trimgalore (Krueger 2015). A montagem do transcriptoma foi feita pelo software Trinity (Grabherr et al. 2012), que computa o melhor k-mer e apresenta alto desempenho para montagem de *contigs* com dados de RNA-seq.

A abundância dos *reads* foi estimada pelo software Kallisto (Bray et al. 2016). A expressão diferencial dos transcritos foi estimada pelo programa EdgeR (Robinson et al. 2009). Posteriormente estes transcritos tiveram a identificação de suas possíveis funcionalidades e atuações no processo celular.

Sequenciamento “long reads” pela tecnologia Oxford Nanopore

A extração do DNA genômico partiu de 30.000 organismos *S. leucops*, e implementado conforme o protocolo fenol/clorofórmio (Sassi et al. 2005). A concentração final de 94 ng/μl de DNA, em 59 μl de água para construção das bibliotecas. As razões 260/280= 1.84 e 260/230= 1.83 pelo Nanodrop. As amostras foram secas em vácuo, na temperatura ambiente em tubos GenTegra™. O sequenciamento foi realizado em plataforma PromethION-Oxford Nanopore pela empresa GenOne, RJ, Brasil.

Os *reads* foram analisados no Fasqc (Andrews 2010), verificando qualidade acima de 20, o que é o recomendado para sequenciamento Nanopore. A montagem foi feita utilizando o software Flye (Kolmogorov 2021), montador *de novo* especializado para seqüências Nanopore, que já faz seu polimento dos *reads*, seleção dos *k-mers* e gera *contigs* curados.

Anotações genoma e transcriptoma

A anotação das regiões codificantes foram feitas no software Augustus (Stanke e Waack 2003), e para o treinamento do software empregamos os genomas de *Macrostomum lignano* (GCA_002269645.1/GCA_001188465.1), *Caenorhabditis elegans* (NC_003279.8/NC_003280.10/NC_003281.10/NC_003282.8/NC_003283.11/NC_003284.9/NC_001328.1) e *Schistosoma mansoni* (NC_031495.1/NC_031496.1/NC_031497.1/NC_031498.1/NC_031499.1/NC_031500.1/NC_031501.1/NC_031502.1/NC_002545.1). Estes genomas foram selecionados pela sua alta cobertura, anotação através de inúmeros métodos, validação de genes, e por serem os mais próximos filogeneticamente a *S. leucops*. Também foi realizada uma análise da completude da montagem empregando o software Busco (Simão et al. 2015) com o banco metazoa_odb10 com 954 proteínas fundamentais. A busca por similaridade foi feita com o banco *seqref* NCBI (<https://www.ncbi.nlm.nih.gov/refseq/>). A montagem das famílias TEs foi feita *de novo* no software RepeatModeler 2.0.2a (Flynn et al. 2020), e a busca por similaridade no RepeatMasker (Smit et al. 2013).

Softwares mais focados na busca de similaridades específicas também foram/estão sendo utilizados:

*Superfamily (Gough et al. 2001) para buscar regiões que possuem similaridade com as principais superfamílias de genes codificantes.

*SignalP3.0 (Bendtsen et al. 2004) para estipular possíveis regiões com *score* que sugeriram a presença proteínas funcionais.

*Promoter2 (Knudsen 1999) para validação de promotores eucarióticos, por similaridade dos já descritos em grande parte dos organismos metazoários.

*WoLF (Horton et al. 2007) para estipular as possíveis áreas de atuação na célula.

Resultados preliminares (breve descrição)

Sequenciamento genoma pela metodologia Illumina

Através do sequenciamento obtido na plataforma Illumina obtivemos 13.643.983 *reads* com tamanhos variando entre 35-301 nucleotídeos. Após o filtro de qualidade 30, obtivemos 10.583.723 *reads* com variação de tamanho entre 50-300 nucleotídeos. Estes *reads* foram usados para a montagem do genoma. As montagens apresentaram os seguintes valores em suas extensões de *contigs*: K-60:n50= 2205, K-45:n50= 2190, K-30:n50=1983, K-21:n50=914. Estes valores são comparáveis a genomas de referência como *Macrostomum lignano*, *Fasciola hepatica* e *Caenorhabditis elegans*. Dentre os diferentes softwares e *K-mers* de montagem o AbySS apresentou os maiores *contigs*, com similaridade mais próxima dos organismos referências. Na montagem, somente 15% dos *contigs* ultrapassaram 700 pb, pois há muitas sequências correspondendo a microssatélites e elementos transponíveis, ultrapassando 40% do conteúdo dos *reads*. As regiões CDSs não expandiram mais de 200 a.a, não possibilitando uma anotação segura. Dos *contigs* gerados 70 % não ultrapassaram 150 a.a de extensão, resultando em 68% como proteínas fragmentadas no Busco e somente 12% completas (14,8% não encontradas e 5,2% duplicadas). Em conjunto, esta montagem se mostrou muito fragmentada, formando muitos *contigs* curtos. Uma opção, para contornar este problema, seria novo sequenciamento Illumina para aumentar a cobertura. Porém, como a análise da montagem obtida apontava para uma fração significativa do genoma, e com elementos transponíveis, optamos por complementar esta montagem com um sequenciamento de *long reads*.

Sequenciamento e montagem do transcriptoma

O sequenciamento de RNA-Seq de vermes com um e dois zooides geraram um total de 80.831.095 *reads*, apresentando tamanhos variando de 40-350 pb. A montagem através do software Trinity gerou 279.881 *contigs*, com n50= 849. Os 279.881 *contigs* gerados pela montagem do Trinity são formados por 64,98% dos *reads* pertencentes a regiões como possíveis CDSs. O software Busco aponta para uma cobertura satisfatória das regiões CDS, onde estão 79,86%-completos, 13,7% fragmentados e 9,44% não encontrados. Quando comparados a organismos referências que contam com dados genômicos disponíveis, vemos: *Macrostomum lignano* (74,9%-completo, 11,6%-fragmentado e 13,5 não encontrados), *Fasciola hepática* (53,8%-completos,

14,8%- fragmentados e 31,4% não encontrado) e *Schmidtea mediterrânea* (70,6%-completo, 9%-fragmentado e 20,4% não encontrado).

Análises com os 279.881 *contigs* gerados através da montagem do Trinity, apenas 489 *contigs* apresentam ser diferencialmente expresso. Dos *contigs* diferencialmente expressos 187 são *up* regulados em I_zooide e 302 são *up* regulados para II_zooides (Figura 1.6). Os *contigs* diferencialmente expressos, 70% destes apresentam alguma região com mascaramento de regiões microsatélites ou TEs.

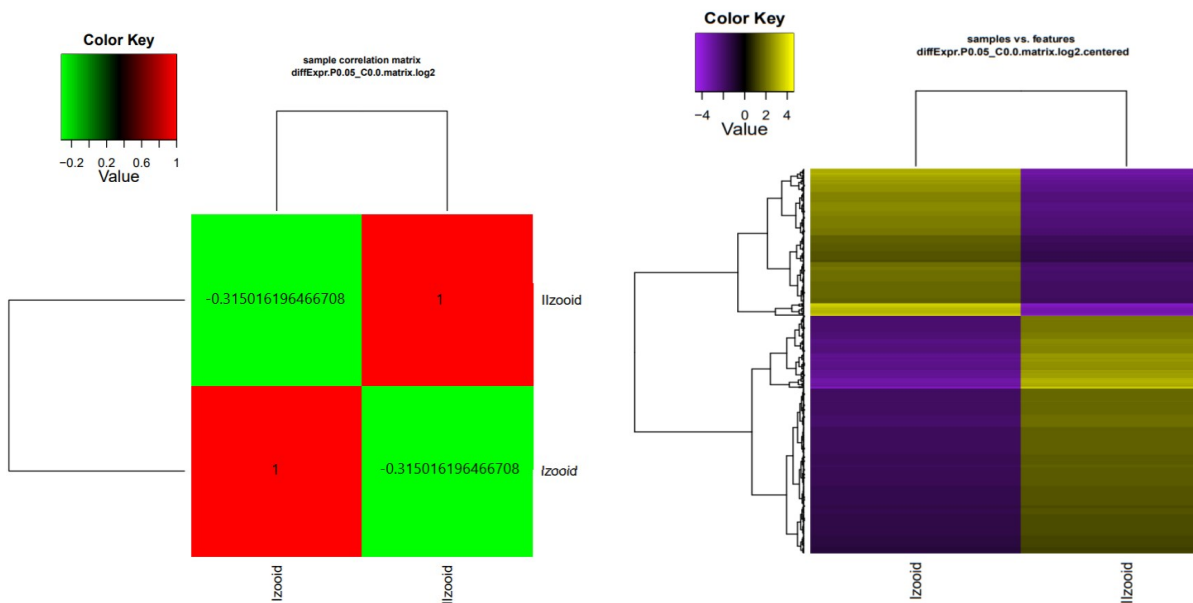


Figura 1.6: Hitmap dos *Contigs* com expressão diferencial significativa entre I_zooide e II_zooides, seguido dos valores de log2 entre as amostras.

Anotações totais foram feitas com todos os transcritos (material suplementar 1.6), mas não sendo correlacionado suas expressões entre os tratamentos. Futuramente será realizada a comparação de expressão com as tabelas de anotação, para verificar os RNAs atuantes em I zooide e II zooide.

Sequenciamento Nanopore

O sequenciamento Nanopore gerou 433.104 *reads*, com tamanhos 112-87.259 pb. Nenhum *read* foi eliminado da montagem, apenas removido as bases finais (50 bases finais) por apresentarem menor qualidade. A cobertura de montagem média ficou em 5,5 vezes, com 5.038 *contigs*. Destes 3.236 apresentaram mais de 10Kb de extensão, com N50=220.186pb e maior com 1.164.702 pb. O tamanho do genoma ficou estipulado em 296.492.128pb, equivalente a 0,29pg de

apresentam *splicing* alternativos, apenas 210 são diferentes para a condição um zoóide e dois zoóides (Figura 4.6).

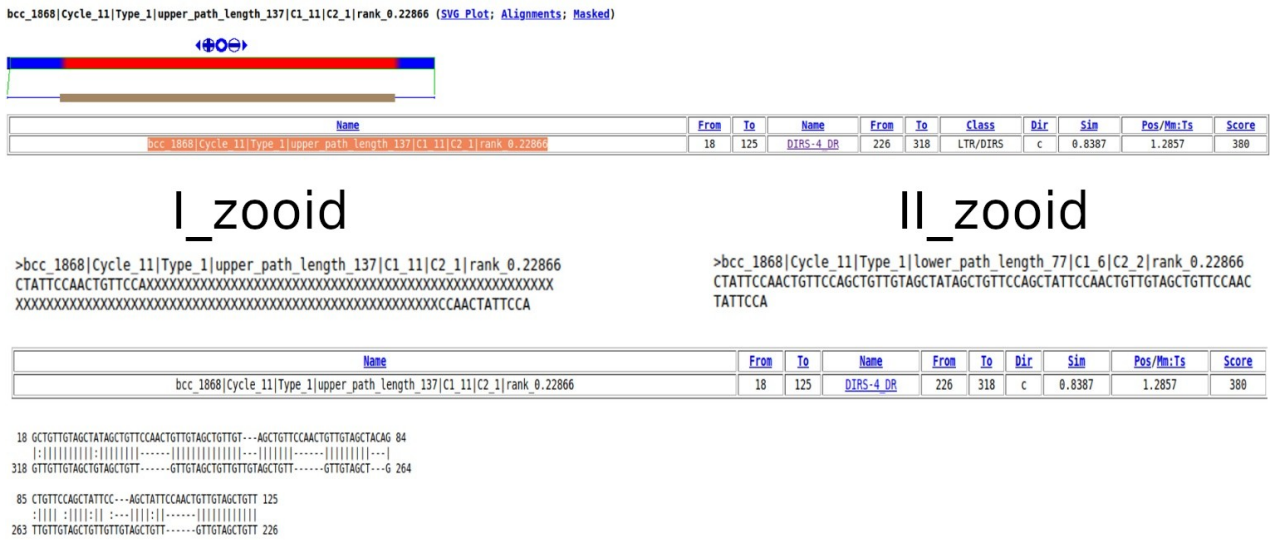


Figura 3.6: Exemplo de *reads* sequenciados com presença de variação, apontados como *splicing*, onde uma de suas variantes possui região com similaridade de elemento transponível (I_zooid) e outra sem mascaramento pela ausência do elemento transponível (II_zooid).

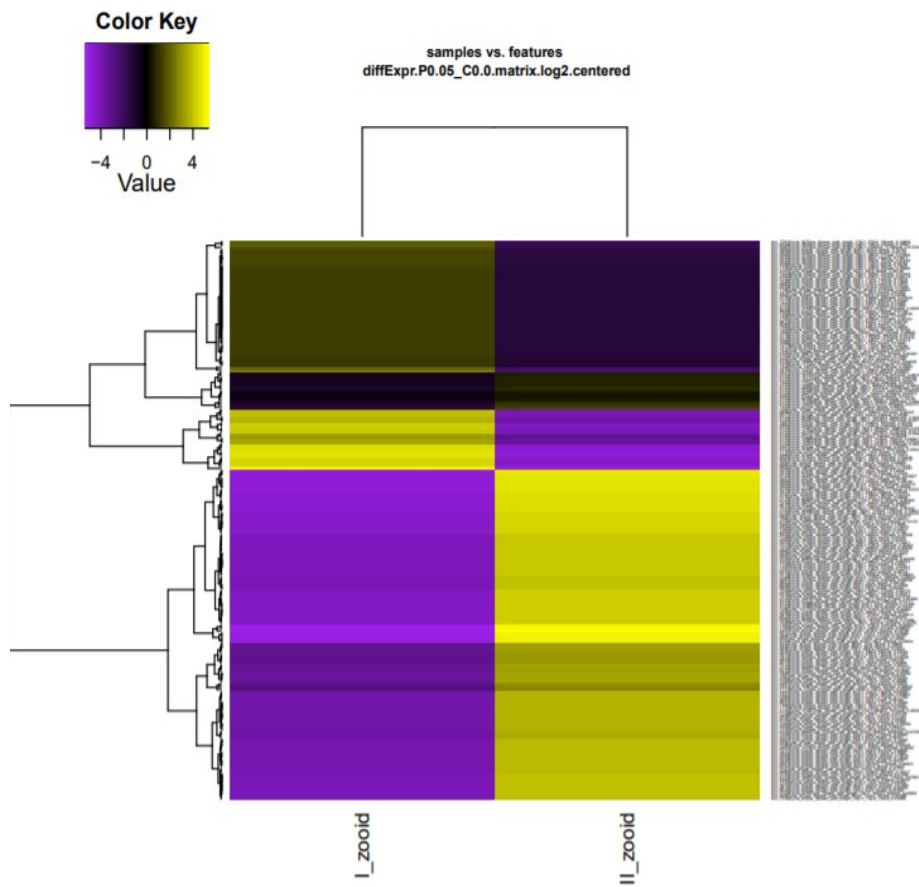


Figura 4.6: Expressão diferencial dos *reads* com *splicing* alternativo, apenas 210 estão sendo diferentemente expresso entre I_zoóide e II_zoóide.

O sequenciamento Nanopore processado pelo RepeaModeler, traz 1650 famílias de TEs. Os *contigs* gerados pelo RepeatModeler 21,11% do seu conteúdo foi mascarado pelo RepeatMasker. As regiões mascaradas estão fragmentadas ao percorrer dos *contigs*, sendo assim necessário uma curagem mais detalhada.

Discussão

Os dados gerados no sequenciamento Illumina não foram capazes de ser informativos sobre as CDSs, pois mais de 40% dos *reads* usados são de regiões repetitivas. As sequências repetitivas nos genomas limitam o sequenciamento, montagem e anotações automatizadas de proteínas (Tørresen et al. 2019). O número elevado de regiões repetitivas nos apontou como possível solução o sequenciamento *long read*, para que pudéssemos contornar esta limitação. O sequenciamento *long read* apresentou uma expansão considerável, saindo de um N50=914 pb (Illumina) para N50=220.186 pb. As famílias de sequências repetitivas geradas pelo RepeatModeler (1650), precisam ser analisadas com mais cuidado, para compararmos com os demais sequenciamentos. A avaliação com relação à montagem do sequenciamento Nanopore feita através do software BUSCO apontou que 53,6 % das ORFs estão completas; 4,8 % duplicadas, 12,8 % fragmentadas e 28,7 % não foram encontradas. Estas estimativas são boas quando comparadas as obtidas para a *Macrostomum lignano* (PacBio + Illumina X 100+ de cobertura) (16,8 completas, 55,2 duplicadas, 11,8 fragmentadas e 16,2 não encontradas). A anotação automática do Augustus sugere 47.839 possíveis proteínas, número próximo ao descrito para *M. lignano* 49.018 CDSs (<https://www.ncbi.nlm.nih.gov/genome/?term=Macrostomum%20lignano>).

Os dados de sequenciamento genômico Nanopore mesmo em baixa cobertura (5,5 X média), sugere que temos uma boa extensão genômica, se compararmos com os pesos moleculares de outros organismos da família Stenostomidae. *S. leucops* ficou estipulado em 0,29 pg, dentro do padrão do peso molecular de citometria para demais organismos da família (*Stenostomum arevaloi* \cong 0.31, *Stenostomum beauchampi* \cong 0.23, *Stenostomum unicolor* \cong 0.20) (<https://www.genomesize.com/>). Para análises futuras estamos pretendendo uma montagem conjunta Illumina e Nanopore. Esta abordagem permite melhorias significativas nas montagens (Chen et al. 2020), assim podemos tentar melhorar nossa cobertura. Análises mais cuidadosas devem ser feitas, pois dados publicados que sugeriram a ausência de genes para *S. leucops* foram encontrados em nossas buscas. Por exemplo, trabalhos evolutivos sobre sistema de silenciamento (piwi, argonauta e siRNA) em Platyhelminthes, apontaram para ausência de genes e RNAs de interferência do complexo Eri (GLD-2) e da via de piRNA (MUT7/NBR) (Fontenla et al. 2021). Entretanto, em nossa análise

preliminar estes genes estão presentes na montagem do genoma de *S. leucops* por nós realizada, com 67,29% em 320 a.a e 92,69% de similaridade em 904 a.a respectivamente. Após uma montagem mais completa, as análises automáticas serão refeitas, além a curagem manual, assim podemos extrair dados mais informativos e cuidadosos.

O transcriptoma também necessitam de uma análise mais cuidadosa, pois a montagem total (I_zooid + II_zooid), e posterior anotação gerou 95.426 candidatos a possíveis regiões CDSs. Destas possíveis CDS, 66,36% têm fragmentos mascarados como repetitivos (TEs e microssatélites), um padrão não convencional comparado com os outros relatos obtidos na literatura. Valores elevados de elementos transponíveis dentro de transcritos gênicos (quimeras), não é algo muito comum. Os dados do Uniprot trazem Invertebrata com 17,34% das CDSs com presença de regiões repetitivas entre transcritos (Tørresen et al. 2019). O trabalho de Christoph e Waddell 2019 encontrou \cong 20% do RNAm total, mas destes 90% destes em regiões de éxons e não são funcionais. As quimeras de RNAm contendo regiões codificantes com TEs tendem a ser deletérias (Lipatov et al. 2005). A taxa considerada neutra ou benéfica em quimeras TEs+CDSs em *Drosophila melanogaster* é 1% (Lipatov et al. 2005). Já o *Homo sapiens* 17% das regiões CDSs têm interação com o retrotransposon L1 (Brandt et al. 2005).

Os transcritos totais de *S. leucops* apresentou 30,65 % de seu conteúdo total mascarado, comparado *Drosophila melanogaster* com 5,66%, empregando a mesma metodologia, sendo 5,4 vezes maior no Platyhelminthe. O Platyhelminthe *M. lignano* apresenta \cong 5% dos seus transcritos mascarados como possíveis TEs (Wasik et al. 2015). As regiões intergênicas em *Macrostomum lignano* são fundamentais para os processos biológicos, manutenção dos neoblastos, e sua diferenciação para os demais tecidos (Azlan et al. 2019). Trabalhos também apontam para efeito benéfico a atividade de elementos transponíveis, como a manutenção de regiões teloméricas no organismo *Drosophila melanogaster*, retardando envelhecimento e manutenção da capacidade reprodutiva celular (Pardue e DeBaryshe 2003). As regiões teloméricas em Platyhelminthes é extremamente importante, para manutenção dos neoblastos (imortalidade dos assexuados eucarióticos).

O trabalho de Blast et al. (2016), traz alguns exemplos, em que organismos de reprodução assexuada tendem a redução de elementos transponíveis em seus genomas. A alta atividade dos TES, pode levar estes organismos a morte, e sendo eliminados pela seleção purificadora. Dados envolvendo outros estudos com organismos como hidras apontam que os elementos transponíveis contribuem para expansão e diversificação dos genomas destes organismos (Wong et al. 2019). Os nossos dados podem ser interessantes para contribuir nestas análises, além de aprofundar o conhecimento das interações moleculares do organismo *S. leucops* e seus elementos transponíveis.

Discussão Geral

Os Platyhelminthes são um dos grupos mais antigos na terra, com a maior diversidade adaptativa dos animais (água doce, água salgada, terrícola e parasitas internos, ou externos) (Sluys 2019). Os organismos intitulados como “organismos modelos” cobrem uma baixa diversidade de organismos, correspondendo a 12 espécies, os quais são os principais alvos para estudos moleculares e celulares. Entretanto, o grande volume de pesquisa em organismos modelo promove uma falta de conhecimento sobre as demais oito a dez milhões de espécies que não são “modelo” (Sullivan 2015). Dentre Platyhelminthes não há um consenso de organismo modelo e que seja amplamente adotado por toda comunidade acadêmica, mas entre os mais estudados estão *Schmidtea mediterranea* (regeneração), *Macrostomum lignano* (modelo emergente), e o parasita *Mesocostoides corti* (cultivo em laboratório). *Stenostomum leucops* pode ser facilmente cultivado em laboratório, produz novo indivíduo por paratomia em +/-48 horas (Rosa et al. 2015). As características citadas anteriormente, com os aspectos apresentados durante a tese apontam *S. leucops* como um excelente organismo experimental. O intuito da tese foi trazer o máximo possível da descrição em relação a biologia de *Stenostomum leucops*, com isto não possuindo um foco único, e muitos destes aspectos podem ser usados em estudos comparativos com outros Catenulida assim como com outros organismos.

O clado dos Catenulidas é representado por organismos de poucos milímetros e com poucas estruturas diferenciadas (Larsson et al. 2008), o que promove dificuldade na classificação taxonômica de seus representantes. A pouca diferenciação celular observada do grupo, faz estes terem predominância de células neoblastos na composição de seus corpos (promovendo a regeneração e reprodução). Os aspectos da biologia básica deste grupo podem ser importantes para entendermos mais da evolução dos Platyhelminthes, e demais filos, já que estes apresentam mais plesiomorfias com os demais invertebrados, que com seus irmãos de clado (Rhabdtophora, Cestoda, Trematoda e Monogênea). Catenulida apresenta aspectos muito divergentes em relação aos demais Platyhelminthes, características de seu desenvolvimento (reprodução, gametogênese, embriogênese), e aspectos moleculares (código genético mitocondrial). Além de ser um grupo pouco estudado, alguns estudos acabam por ser contraditórios, um exemplo era sua regeneração (capítulo 3), e ainda persiste dúvidas se os representantes dessa classe é de reprodução assexuada obrigatória ou se eventuais casos de reprodução sexuada ocorrem (estamos começando alguns estudos neste sentido).

A posição filogenética de Catenulida está em discussão até os dias atuais. Os trabalhos mais clássicos levaram em conta somente os aspectos morfológicos, assim Catenulida era considerado

basal ao grande grupo dos vermes achatados (Platyhelminthes, Acoela e Nemertodermatida) (Karling 1974, Ehlers 1985, Haszprunar 1996). Os artigos mais atuais, empregando informações de biologia molecular e biologia do desenvolvimento, removeram Acoela e Nemertodermatida de Platyhelminthes, e agrupam Catenulida como irmão de Rhabditophora (Carranza et al. 1997, Litvaitis e Rohde 1999, Telford et al. 2003, Wallberg et al. 2007 Egger et al. 2015, Laumer et al. 2015). O capítulo 2 apresenta a análise filogenética de 12 CDSs com 165 mitogenomas de Platyhelminthes, reforçando Catenulida como grupo irmão de Rhabditophora. Além de ressaltar os aspectos peculiares do genoma mitocondrial de Catenulida, como a presença de CDSs nas duas fitas do DNA mitocondrial, ordem gênica distinta, codificação tRNA-K plesiomórfico aos demais invertebrados não Platyhelminthes. Além disso, o sequenciamento Nanopore sugere que o mitogenoma apresenta uma grande duplicação, sendo o mitogenoma maior que o obtido pela montagem de *short reads*. Estes dados são concordantes com outros achados, para outros organismos, de que montagem de sequenciamento de *short reads* pode subestimar o tamanho dos mitogenomas.

A biologia molecular de *S. leucops* apresenta características não comuns, que vão além da mitocondrial, como a utilização de DNA presente no ambiente e/ou mantido em seu microbioma. O capítulo 3 nós trazemos dados moleculares robustos sobre a interação molecular *S. leucops* + microbioma. O Platyhelminthe Catenulida *Paracatenula sp.* apresenta uma troca direta de metabolitos com seu endossimbionte *Candidatus Riegeria galateiae*, que ainda não haja indícios e troca de moléculas de DNA entre eles (Gruber-Vodicka et al. 2011, Jäckle et al. 2018). *S. leucops* não apresenta uma interação de dependência tão grande como *Paracatenula sp.*, pois há uma variação bem mais ampla em seu microbioma no tempo (decorrer de anos de cultivo) e espaço (local de cultivo), estes aspectos foram trazidos no Capítulo 5. Na tentativa de salvar a cultura de *S. leucops* em um *bottleneck*, fizemos um enriquecimento significativo da diversidade do microbioma. A causa principal do *bottleneck*, nos leva a alteração nas condições da água do cultivo, pois apresenta a bactéria *Rhodospirillum rubrum* redutora de Fe³⁺, com alta capacidade de crescimento (Finneran et al. 2003). A linhagem SL0-sm02 apresenta 42,7% das sequências pertencerem a *Rhodospirillum rubrum*. Além da bactéria ainda se manter em SL0-sm03, os índices de ferro atual estão altos (813µg/L). Os valores de diversidade alfa saltaram de 0,43 no princípio da cultura (SL0-sm01), para 3,05 com a inserção de outras fontes de água (SL0-sm02). Após a estabilização do cultivo novamente em laboratório, o índice de diversidade decaiu para 1,36 (SL0-sm03), o mesmo método de cultivo, com outra fonte de água, está em 2,16 (SL0-hc). Atualmente os dois microbiomas vigentes de cultivo (SL0-sm03 e SL0-hc), junto com organismos encontrados em ambientes naturais (SL0-wp), mesmo com índices de diversidade distintos a

predominância está em quatro gêneros: *Bacillus*, *Pseudomonas*, *Flectobacillus* e *Ralstonia* (96-98% dos *reads*). O gênero *Bacillus* é abundantemente nos quatro cultivos com facilidade de manutenção de *S. leucops*, e somente não sendo expressivamente abundante no *bottleneck*. Estudos mais detalhados precisam ser feitos, para verificar se há uma interação mais específica *S. leucops* e *Bacillus*. Verificamos que genes contidos em plasmídeos mantidos por *E. coli* (ou *E. fergusonii*) podem ser expressos por *S. leucops*. Alguns resultados prévios sugerem que *Bacillus* também podem manter estes plasmídeos indutores de expressão de *gfp*.

Os avanços da tecnologia de sequenciamento permitem gerar uma quantidade imensa de dados. Porém sem uma anotação cuidadosa do genoma montado não se avança na compreensão sobre o funcionamento desses genomas e sobre a evolução genômica (Sullivan 2015). O filo Platyhelminthe apresenta, no momento, 48 genomas sequenciados (quatro de vida livre e 44 parasitas), desses, apenas 23 sequenciamentos são anotados com suas proteínas e demais constituintes genômicos. *M. lignano* é o único representante de vida livre com o genoma anotado. A anotação automática da montagem do sequenciamento Nanopore que realizamos, apresentou 47.839 possíveis CDSs, próximo aos 49.018 descritos para *M. lignano*. No entanto, outras análises comparativas ainda precisam ser feitas. O genoma de *M. lignano* apresenta 75% do seu conteúdo como sequências repetitivas (TEs, microssatélites) (Wasik et al. 2015). As análises de TEs de *S. leucops* também precisam ser mais refinada, pois 90% dos fragmentos apontados como elementos de transposição, estão dispostos em fragmentos curtos. Os fragmentos curtos podem ser realmente indicadores de acúmulos de TEs degenerados, ou único elemento, assim precisam de curagem manual e descrição. O sistema de silenciamento em *M. lignano* tem como principal função a manutenção dos neoblastos (Palakodeti et al. 2008, Zhou et al. 2015). *M. lignano* possui reprodução sexuada, e seus neoblastos têm principal função no crescimento dos tecidos e regeneração quando necessário. Já *S. leucops* depende ativamente das células-tronco (*stem cells*), para reprodução e regeneração.

A “possível falhas no silenciamento” de elementos transponíveis, pode ser um promotor da plasticidade, diversidade em organismos assexuados (revisão em Loreto e Moura 2017). *S. leucops* apresenta 36 dos 54 genes envolvidos no silenciamento de TEs pelas vias de pequenos RNAs. *M. lignano* possui 43 segundo Fontenla et al. (2021). Em nossas análises prévias encontramos mais duas possíveis (Eri complexo “GLD-2” e via de piRNA “MUT7/NBR”). A alta prevalência de sequências repetitivas (30,65 % do conteúdo) nos mRNAs, causa questionamentos. Estas sequências teriam funcionalidade ou seriam apenas inserções acidentais, neutras, que nada causam no funcionamento desses genomas? A resposta à pergunta sobre a função de TEs transcritos junto às CDSs precisa de uma análise mais cuidadosa, já que a literatura não traz isto como algo “comum”. Análises com aprendizado de máquina (*machine learning*) poderão ajudar neste trabalho, e quem

sabe serem aplicadas aos demais genomas não anotados. Os trabalhos de biologia básica também devem prosseguir, para assim aprendermos mais sobre os muitos segredos dos Platyhelminthes. Para no futuro entendermos sobre o sucesso evolutivo de um dos primeiros grupo de organismos a viverem na terra e perdurarem até hoje, nos mais diversos habitats e modos de vida.

Referências:

- Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Azlan, A., Halim, M. A., Azzam, G. (2019). Genome-wide identification and characterization of long intergenic noncoding RNAs in the regenerative flatworm *Macrostomum lignano*. *Genomics*. doi:10.1016/j.ygeno.2019.07.016
- Bast, J., Schaefer, I., Schwander, T., Maraun, M., Scheu, S., Kraaijeveld, K. (2016). No Accumulation of Transposable Elements in Asexual Arthropods. *Molecular biology and evolution*, V. 33, pp 697–706. <https://doi.org/10.1093/molbev/msv261>
- Bordenstein, S. R. and Theis, K. R. (2015). Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biol.* V. 13, e1002226.
- Borkott, H. (1970). Geschlechtliche Organisation, Fortpflanzungsverhalten und Ursachen der sexuellen Vermehrung von *Stenostomum sthenum* nov. spec. (Turbellaria, Catenulida). *Z. Morph. Tiere* V. 67, pp 183-262.
- Chen, Z., Erickson, D.L. and Meng, J. (2020). Benchmarking hybrid assembly approaches for genomic analyses of bacterial pathogens using Illumina and Oxford Nanopore sequencing. *BMC Genomics* V. 21, pp 631. <https://doi.org/10.1186/s12864-020-07041-8>.
- Egger, B., Bachmann, L. and Fromm, B. (2017). Atp8 is in the ground pattern of flatworm mitochondrial genomes. *BMC Genomics* V. 18, pp 414. <https://doi.org/10.1186/s12864-017-3807-2>.
- Ehlers, U. and Sopott-Ehlers, B. (1995). Plathelminthes or Platyhelminthes?. *Hydrobiologia* V. 305, pp 1–2. <https://doi.org/10.1007/BF00036354>.
- Fontenla, S., Rinaldi, G., Tort, J. F. (2021). Lost and Found: Piwi and Argonaute Pathways in Flatworms.- *Frontiers in Cellular and Infection Microbiology*. V. 11 pp 2235-2988. <https://www.frontiersin.org/article/10.3389/fcimb.2021.653695>.

Gruber-Vodicka, H. R., Dirks, U., Leisch, N., Baranyi, C., Stoecker, K., Bulgheresi, S., Heindl, N. R., Horn, M., Lott, C., Loy, A., Wagner, M., Ott, J. (2011). *Paracatenula*, an ancient symbiosis between thiotrophic *Alphaproteobacteria* and catenulid flatworms. *Proceedings of the National Academy of Sciences*. V. 29 pp 12078-12083. DOI: 10.1073/pnas.1105347108.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*, V. 117(17), pp 9451–9457. <https://doi.org/10.1073/pnas.1921046117>.

Haag, K. L. (2018). Holobionts and their hologenomes: Evolution with mixed modes of inheritance. *Genet Mol Biol*. V. 41, pp 189–197.

Hahn, C., Fromm, B., Bachmann, L. (2014). Comparative genomics of flatworms (platyhelminthes) reveals shared genomic features of ecto-and endoparasitic neodermata. *Genome biology and evolution*. V. 6(5), pp 1105–1117. <https://doi.org/10.1093/gbe/evu078>.

Hayes, M. J. (2017). Sulphated glycosaminoglycans support an assortment of planarian rhabdite structures. *Biol*. V. 6, pp. 571-581.

Gruber-Vodicka, H. R., Dirks, U., Leisch, N., Baranyi, C., Stoecker, K., Bulgheresi, S., Heindl, N. R., Horn, M., Lott, C., Loy, A., Wagner, M., Ott, J. (2011). *Paracatenula*, an ancient symbiosis between thiotrophic *Alphaproteobacteria* and catenulid flatworms. *PNAS*. V. 108 pp 12078-12083. <https://doi.org/10.1073/pnas.1105347108>

Jäckle, O. et al. (2019). Chemosynthetic symbiont with a drastically reduced genome serves as primary energy storage in the marine flatworm *Paracatenula*. *Proc. Natl. Acad. Sci*. V. 116 pp 8505-8514.

Lee, F. J., Williams, K. B., Levin, M., Wolfe, B. E. (2018). The Bacterial Metabolite Indole Inhibits Regeneration of the Planarian Flatworm *Dugesia japonica*. *Iscience*, V. 10, pp 135-148. <https://doi.org/10.1016/j.isci.2018.11.021>.

Lipatov, M., Lenkov, K., Petrov, D. A., Bergman, C. M. (2005). Paucity of chimeric gene-transposable element transcripts in the *Drosophila melanogaster* genome. *BMC biology* V. 24. <https://doi.org/10.1186/1741-7007-3-24>

Kolmogorov, M. (2021). *Flye*. *GitHub repository*. GitHub. Retrieved from <https://github.com/fenderglass/Flye>.

Smit, A., Hubley, R., Green, P. (2013). *RepeatMasker Open-4.0*. <http://www.repeatmasker.org/webrepeatmaskerhelp.html>.

Sluys, R. (2019). The evolutionary terrestrialization of planarian flatworms (Platyhelminthes, Tricladida, Geoplanidae): a review and research programme. *Zoosystematics and Evolution*. V. 95 pp 543-556. <https://doi.org/10.3897/zse.95.38727>.

Pardue, M. L. and DeBaryshe, P. G. (2003). Retrotransposons provide an evolutionarily robust non-telomerase mechanism to maintain telomeres. *Annu Rev Genet*. V. 37, pp 485-511. doi: 10.1146/annurev.genet.38.072902.093115. PMID: 14616071.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, V. 31(19), pp 3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>.

Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*, V. 19 pp 215–225. <https://doi.org/10.1093/bioinformatics/btg1080>.

Sullivan, W. (2015). The Institute for the Study of Non-Model Organisms and other fantasies. *Molecular biology of the cell*. V. 26 pp 387–389. <https://doi.org/10.1091/mbc.E14-03-0814>

Tørresen, O. K. et al. (2019). “Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases.” *Nucleic acids research*. V. 47,21, pp 10994-11006. doi:10.1093/nar/gkz841.

Wasik, K., Gurtowski, J., Zhou, X., Ramos, O.M., Delás, M.J., Battistoni, G., Demerdash, O., Falciatori, I., Vizoso, D.B., Smith, A.D., Ladurner, P., Schärer, L., McCombie, W.R., Hannon, G.J.,

Schatz, M. (2015). *Macrostomum lignano* genome. Proceedings of the National Academy of Sciences, V. 112 (40) pp 12462-12467; DOI: 10.1073/pnas.1516718112.

Wicker, T., Sabot, F., Hua-Van, A. *et al.* (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* V. 8 pp 973–982. <https://doi.org/10.1038/nrg2165>

Wong, W. Y., Simakov, O., Bridge, D. M., Cartwright, P. , Bellantuono, A. J., Kuhn, A., Holstein, T. W., David, C. N., Steele, R. E., Martínez, D. E. (2019). Expansion of a single transposable element family is associated with genome-size increase and radiation in the genus *Hydra*. *PNAS* V. 46 pp 22915-22917. <https://doi.org/10.1073/pnas.1910106116>.

Material Suplementar:

*Material supl. 1.6: Tabela anotação prévia dos *contigs* transcriptoma. (https://www.dropbox.com/scl/fi/9hn1jd5w2qwgjwjkubs7k/sten_transcriptoma.ods?dl=0&rlkey=lor8hf034pes2lk581gzs3mlq).

*Material supl. 2.6: Tabela anotação prévia dos *contigs* Nanopore. (https://www.dropbox.com/scl/fi/4j96oyv19f0tgud89eenm/dados_totais_nanopore_sten.xls?dl=0&rlkey=t9mti418o13t5qzsq8frp8li).