

UFRRJ

INSTITUTO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE

**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA
ANIMAL**

TESE

**Evolução de DNA Satélite na ordem Acipenseriformes
(Vertebrata: Actinopterygii): Uma abordagem genômica**

JOSÉ HENRIQUE FORTE STORNIOLI

2025



UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL

**Evolução de DNA Satélite na ordem Acipenseriformes (Vertebrata:
Actinopterygii): Uma abordagem genômica**

JOSÉ HENRIQUE FORTE STORNIOLI

Sob a Orientação do Professor Doutor

Ricardo Utsunomia

e Co-orientação do Professor Doutor

Fabio Porto-Foresti

Tese submetida como requisito para obtenção do grau de **Doutor em Biologia Animal**, no curso de Pós-Graduação em Biologia Animal, Área de Concentração em Biodiversidade Animal.

Seropédica, RJ

Dezembro de 2025

Ficha Catalográfica

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

S863e Stornioli, José Henrique Forte, 1993-
Evolução de DNA Satélite na ordem Acipenseriformes
(Vertebrata: Actinopterygii): Uma abordagem genômica /
José Henrique Forte Stornioli. - São Paulo, 2025.
64 f.

Orientador: Ricardo Utsunomia.
Coorientador: Fabio Porto-Foresti.
Tese(Doutorado). -- Universidade Federal Rural do
Rio de Janeiro, Biologia Animal, 2025.

1. Satelitômica. 2. Evolução em Concerto. 3.
Hipótese Library. I. Utsunomia, Ricardo, 1987-,
orient. II. Porto-Foresti, Fabio, 1972-, coorient.
III Universidade Federal Rural do Rio de Janeiro.
Biologia Animal. IV. Título.

O presente trabalho foi realizado com apoio da **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)**, Código de Financiamento 001. This study was financed in part by the **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)** - Finance Code 001.



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL



TERMO Nº 1212 / 2025 - PPGBA (12.28.01.00.00.00.42)

Nº do Protocolo: 23083.074057/2025-34

Seropédica-RJ, 19 de dezembro de 2025.

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
CURSO DE PÓS-GRADUAÇÃO EM BIOLOGIA ANIMAL

JOSE HENRIQUE FORTE STORNIOLI

Tese submetida como requisito parcial para obtenção do grau de
Doutor(a) em Ciências, no Curso de pós-graduação em **BIOLOGIA ANIMAL**,
área de concentração em **BIODIVERSIDADE ANIMAL**

DISSERTAÇÃO APROVADA EM 16/12/2025

(Assinado digitalmente em 22/12/2025 20:50)

DENISE MONNERAT NOGUEIRA
PROFESSOR DO MAGISTERIO SUPERIOR
DeptG (12.28.01.00.00.00.50)
Matrícula: 1539011

(Assinado digitalmente em 30/12/2025 08:36)

RICARDO UTSUNOMIA
ASSINANTE EXTERNO
CPF: 362.387.008-07

(Assinado digitalmente em 19/12/2025 10:51)

SERGIO PEREIRA
ASSINANTE EXTERNO
CPF: 277.399.348-95

(Assinado digitalmente em 04/02/2026 15:36)

MARCELO DE BELLO CIOFFI
ASSINANTE EXTERNO
CPF: 073.339.076-51

(Assinado digitalmente em 23/12/2025 08:09)

FRANCISCO DE MENEZES CAVALCANTE SASSI
ASSINANTE EXTERNO
CPF: 236.107.998-46

Visualize o documento original em <https://sipac.ufrj.br/public/documentos/index.jsp>
informando seu número: **1212**, ano: **2025**, tipo: **TERMO**, data de emissão: **19/12/2025** e o
código de verificação: **9bae32c78f**

Agradecimentos

Início meus agradecimentos a todos os animais que foram utilizados ao longo dos meus 13 anos desde que iniciei minha jornada no mundo acadêmico, diversos *taxa* foram estudados, sempre manuseados e estudados com o maior respeito possível e respeitando os mais estritos preceitos éticos na experimentação animal. Meus mais sinceros agradecimentos por permitirem o desenvolvimento científico do *Homo sapiens*.

O presente trabalho foi realizado com apoio da **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)**, sob projeto número 88887.799526/2022-00.

Agradeço a todos os servidores e servidoras, terceirizados e demais trabalhadores e trabalhadoras, que permitiram que eu terminasse este trabalho, e meu agradecimento é literalmente a todos, desde os guardas que sempre encontrava, até quando ia trabalhar de final de semana, até àqueles que não tive contato além do e-mail, vocês são responsáveis pela formação de centenas, milhares de estudantes, e devem ser reconhecidos e valorizados, meus mais sinceros agradecimentos por permitir e dar suporte a nós.

Agradeço ao Prof. Dr. Fábio Porto-Foresti e ao Prof. Dr. Ricardo Utsunomia por todo o ensinamento ao longo do meu doutorado, por me ensinar o que é ser cientista, como este mundo funciona, como me desenvolver, aproveitar oportunidades, e a ser uma pessoa melhor. O mundo acadêmico não é fácil, mas os ensinamentos e o que eu vivi foram fundamentais para eu terminar o Doutorado preparado para o que der e vier nas próximas etapas.

A todos os meus colegas do Laboratório de Genômica e Conservação de Peixes – LAGENPE Unesp Bauru por todos esses anos de trabalho, tivemos diversos momentos ótimos, eu aprendi com todos vocês e sei que contribuí para a formação de cada um, agradeço por todos os momentos bons, que também foram fundamentais para finalizar este trabalho.

A palavra agradecimento é pouco para o que minha mãe Rosimeire Forte Stornioli e as minhas irmãs Caroline Forte Stornioli e Victória Forte Stornioli fizeram para que eu pudesse estar aqui hoje, só tenho a agradecer porque sem vocês eu não estaria aqui e não teria esse título, amo vocês e obrigado por me permitir ir atrás deste sonho, que outrora era distante, difícil, praticamente inalcançável, e agora estou aqui.

À minha namorada Larissa Roncon, que me deu suporte e apoio nesta etapa tão fundamental da minha carreira e formação, me ajudando nos momentos difícil, me dando suporte nos momentos que tudo estava deveras complicado, e agora que vai comemorar comigo a felicidade de terminar essa etapa tão difícil.

Aos meus queridos amigos e amiga, quero dizer, aos meus irmãos e irmã que fiz nestes últimos anos, Vitor, Futurinho e Amanda, que acompanharam essa jornada de perto e me deram forças para continuar, apesar de todas as dificuldades, que não foram poucas, todos os momentos ótimos que passamos foram fundamentais para que eu pudesse terminar essa etapa, e espero que vocês saibam disso.

Agradeço também ao Zeni, a Nati, a Gabi e ao Kelvin por todos esses anos de amizade e companheirismo, vocês foram fundamentais para que eu conseguisse chegar onde cheguei, tanto ajudando direta ou indiretamente nos trabalhos, e também com muita diversão, alegrias, discussões e conversas que foram extremamente produtivas, vocês são deveras importantes nesta etapa da minha vida.

À rapaziada do bosque, que estiveram comigo desde antes desta etapa, e que continuaram comigo, mesmo a vida tendo levado muitos a caminhos diferentes, os momentos de alegria ficarão e foram importantes para chegar aonde estou. Borracha virou doutor rapeize...

UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
CAMPUS DE SEROPÉDICA
DEPARTAMENTO DE CIÊNCIAS BIOLÓGICAS

Evolução de DNA Satélite na ordem Acipenseriformes (Vertebrata: Actinopterygii): Uma abordagem genômica

Discente: José Henrique Forte Stornioli

Orientador: Prof. Dr. Ricardo Utsunomia

Coorientador: Prof. Dr. Fábio Porto-Foresti

Tese apresentado à Universidade Federal Rural do Rio de Janeiro, departamento de Biologia Animal, Campus de Seropédica, para a **Defesa de Doutorado**, como requisito para a obtenção do título de Doutor em Biologia Animal.

Seropédica

2025

Resumo

Com a popularização do sequenciamento de nova geração tornou-se possível estudar elementos que, outrora, era inacessível o estudo massivo, como por exemplo os DNAs satélites. Com isso, houve um grande desenvolvimento de formas de se estudar estes elementos e melhor compreender sua possível funcionalidade e dinâmica evolutiva. Portanto, visando descrever os processos evolutivos dessas sequências em *Actionpterygii*, seis catálogos de DNAs satélites foram descritos, com 233 novos DNAs satélites distribuídos nas seis espécies, com 106 DNAs satélites compartilhados entre todas as espécies exceto *Polypterus senegalus*, a maioria deles no gênero *Acipenser*, corroborando a predição da hipótese *Library*. Além disso, foi caracterizado 12 DNAs satélites altamente conservados (>95% de similaridade) e observou-se que um conjunto de satDNAs foi encontrado apenas no gênero *Acipenser*, não ocorrendo nos demais grupos externos, demonstrando que este conjunto de sequências possui uma origem de, pelo menos, 100 ma. Por fim, o satDNA *HindIII*, que já foi isolado em outras espécies de *Acipenser*, foi isolado no presente estudo enquanto outros satDNAs foram encontrados em possíveis regiões de lncRNA ou mRNA de *scaffolds* não ancorados de *Acipenser ruthenus*.

Palavras-chave: Satelitômica comparativa, Evolução em Concerto, Hipótese *Library*.

Abstract

With the popularization of next-generation sequencing, it has become possible to study elements that were previously inaccessible to large-scale study, such as satellite DNAs. This has led to significant developments in methods for studying these elements and better understanding their potential functionality and evolutionary dynamics. Therefore, aiming to describe the evolutionary processes of these sequences in the *Actionpterygii* clade, six catalogs of satellite DNAs were described, with 233 new satellite DNAs distributed across the six species, with 106 satellite DNAs shared among all species except *Polypterus senegalus*, most of them in the genus *Acipenser*, corroborating the prediction of the Library hypothesis. Furthermore, 12 highly conserved satellite DNAs (>95% similarity) were characterized, and it was observed that a set of satDNAs was found only in the genus *Acipenser*, not occurring in other outgroups, demonstrating that this set of sequences has an origin of at least 100 million years ago. Finally, satDNA *HindIII*, which has already been isolated in other *Acipenser* species, was isolated in the present study, while other satDNAs were found in possible lncRNA or mRNA regions of unanchored scaffolds of *Acipenser ruthenus*.

Keywords: Comparative satelitomics, Concerted evolution, Library Hypothesis.

Sumário

1. Introdução.....	1
1.1 Genoma de organismos eucarióticos e DNA satélite	1
1.2 Aspectos genômicos de Acipenseriformes	3
1.3 Importância da ordem Acipenseriformes	4
2.1. Objetivo Geral.....	6
2.2. Objetivos específicos	6
3. Material e métodos.....	7
3.1 Espécies utilizadas	7
3.2 Caracterização dos satelitomas.....	7
3.2 Satelitômica comparativa e BLAST.....	9
4. Resultados.....	11
4.1 Caracterização dos satelitomas.....	11
4.1.1 <i>Polyodon spathula</i>.....	11
4.1.2 <i>Acipenser oxyrinchus</i>.....	12
4.1.3 <i>Acipenser schrenckii</i>	12
4.1.4 <i>Acipenser sinensis</i>	12
4.1.5 <i>Acipenser ruthenus</i>	13
4.1.6 <i>Polypterus senegalus</i>.....	13
4.2 Satelitômica comparativa na ordem Acipenseriformes	13
4.3 Resultados do BLAST.....	22
5. Discussão.....	27
6. Conclusão.....	31
7. Referências.....	31
8. Material Suplementar	38

1. Introdução

1.1 Genoma de organismos eucarióticos e DNA satélite

O genoma de organismos eucarióticos apresenta diversos elementos que conferem sua complexidade única. Entre eles, destacam-se os genes responsáveis pela expressão de proteínas ao longo da vida dos organismos; os pseudogenes, que são cópias de genes que perderam sua função devido a mutações e que podem surgir tanto por duplicação quanto por retrotransposição, influenciando a regulação gênica; e os elementos repetitivos, cuja principal função é a regulação do genoma e de sua estrutura (Biscoti et al., 2015).

Os elementos repetitivos podem ser divididos em duas grandes categorias: i) os elementos repetitivos dispersos e ii) os elementos repetitivos em tandem (Yuan et al., 2018). A primeira categoria inclui transposons e retrotransposons, que possuem a capacidade de se mover pelo genoma e se duplicar, inserindo-se em regiões distintas (Kojima, 2019). Já a segunda categoria é composta por famílias multigênicas e DNAs satélites (satDNAs). As famílias multigênicas são formadas pelo gene e pelo espaçador, que juntos compõem o monômero. Por outro lado, os satDNAs consistem em monômeros organizados em conformação cabeça-cauda, formando longos *arrays* repetidos, que podem alcançar centenas a milhares de cópias e ser encontrados em diferentes regiões do genoma, como as centroméricas, pericentroméricas, intersticiais e subteloméricas, preferencialmente associadas à heterocromatina, sendo um elemento estrutural do genoma (Garrido-Ramos, 2017; Ruiz-Ruano et al., 2016).

Os DNAs satélites (satDNAs) são elementos altamente diversos, cujas características incluem composição de bases, conteúdo A+T, tamanho do monômero, extensão dos *arrays*, que podem ser longos e concentrados em regiões específicas do genoma ou curtos e dispersos e número de cópia no genoma (Šatović-Vukšić; Plohl, 2023). Segundo a teoria da evolução em concerto (Elder; Turner, 1995), os monômeros de satDNA evoluem de maneira coordenada, de modo que mutações ocorridas nessas sequências são homogeneizadas dentro do indivíduo e, posteriormente, na população, ou eliminados. Esse processo ocorre por meio da fixação ou eliminação das mutações nos *arrays* de satDNA, promovido por mecanismos de troca não recíproca, como *crossing-over* desigual e conversão gênica (Elder; Turner, 1995; Šatović-Vukšić; Plohl, 2023; Smith, 1976).

Sob uma perspectiva evolutiva, é possível observar que espécies próximas compartilham essas sequências, embora apresentem padrões distintos de abundância e divergência. Em alguns casos raros, essas sequências podem ser conservadas em grandes grupos taxonômicos, enquanto, em outras situações, novos satDNAs podem surgir independentemente nessas espécies (Dos Santos et al., 2021; Garrido-Ramos, 2017; Ruiz-Ruano et al., 2016; Sales-Oliveira et al., 2024).

Para explicar a diversidade de DNAs satélites entre diferentes espécies, Fry & Salser (1977) propuseram a hipótese *Library*. Esta hipótese afirma que espécies próximas compartilham DNAs satélites, mas com perfis distintos de abundância, enquanto espécies mais distantes podem não compartilhar DNAs satélites, possuindo uma "biblioteca" completamente diferente. Estudos recentes corroboram essa hipótese, demonstrando que espécies próximas, como membros dos gêneros *Psalidodon* e *Astyanax* (Goes et al., 2022), compartilham DNAs satélites, assim como membros da família Serrasalminidae (Goes et al., 2023). Estudos realizados com duas espécies de gafanhotos demonstraram que ambas as espécies possuem bibliotecas específicas de satDNA que são resultados de processos de amplificação e contração específicos oriundos de um conjunto de satDNAs presentes no ancestral em comum (Camacho et al., 2022) e, observando grupos mais amplos, foi identificado um DNA satélite conservado na subordem *Characoidei*, denominado CharSat01-52 (dos Santos et al., 2021), reforçando a validade da hipótese *Library* ao explicar os diferentes padrões de abundância observados tanto em espécies próximas quanto em espécies mais distantes.

O advento do sequenciamento de segunda geração reduziu significativamente os custos para sequenciar organismos não-modelo, e com o desenvolvimento de *pipelines* específicas para caracterização de elementos repetitivos, principalmente DNAs satélites (Ruiz-Ruano et al., 2016). Os estudos com esses elementos ganharam notoriedade e permitiram um maior conhecimento acerca da dinâmica evolutiva, localização genômica, bem como as possíveis funções. Como exemplo, na infraclasse Teleostei, estudos que observaram o acúmulo de DNAs satélites em cromossomos B (Stornioli et al., 2021), em cromossomos sexuais (De Moraes et al., 2023) e estudos comparativos (dos Santos et al., 2021; Utsunomia et al., 2017) demonstraram que é possível aliar a bioinformática com a citogenética clássica, dando origem à citogenômica, área que visa justamente responder grandes perguntas existentes há décadas e levantar novas questões devido à quantidade massiva de dados gerados pela bioinformática e as questões naturalmente desenvolvidas e que serão trabalhadas nos próximos anos.

Nos Actinopterygii basais (Hughes et al., 2018), os estudos com DNAs satélites concentram-se principalmente na ordem Acipenseriformes. Dois DNAs satélites conservados foram descritos por Robles e colaboradores (2004), onde observaram que uma das principais características do grupo, a evolução molecular lenta, reflete-se na evolução desses elementos repetitivos. Os dois DNAs satélites conservados, PstI e *Hind*III, apresentam taxas de mutação de 0,07% a 0,11% por milhão de anos, respectivamente, e o DNA satélite *Hind*III é amplamente distribuído na ordem Acipenseriformes (Lanfredi et al., 2001). Outros estudos utilizaram DNAs satélites para investigar a presença de cromossomos sexuais em *Acipenser ruthenus*, identificando possíveis pares cromossômicos com acúmulo de sequências repetitivas (Biltueva et al., 2017). Além disso, foi descrito que, após eventos de duplicação genômica exclusivos deste grupo, a distribuição cromossômica dos DNAs satélites correlaciona-se com as alterações cariotípicas observadas em *Acipenser baerii* (Biltueva et al., 2020).

1.2 Aspectos genômicos de Acipenseriformes

Ao longo da história evolutiva dos vertebrados, ocorreram dois eventos de duplicação genômica na base do grupo, há aproximadamente 600 milhões de anos, resultando em alterações genômicas significativas ao longo do tempo (Dehal; Boore, 2005). Dentro da infraclasse Teleostei, um terceiro evento de duplicação genômica ocorreu há cerca de 320 milhões de anos (Vandepoele et al., 2004). Esses eventos podem ser vantajosos para as espécies, pois promovem alterações epigenéticas, mudanças na expressão gênica e nas vias de regulação do genoma, potencializando a adaptação a condições ambientais extremas ou, até mesmo, levando ao surgimento de novas espécies (Rieseberg et al., 2007; Sémon; Wolfe, 2007; Van De Peer; Maere; Meyer, 2009).

Em Actinopterygii, grande parte dos *taxa* que compõe o grupo não apresentam evidências de eventos de duplicação genômica, exceto na ordem Acipenseriformes, que apresentam eventos de duplicação exclusivos, resultando em números diplóides elevados, superiores a 200 cromossomos em algumas espécies, refletindo um elevado grau de poliploidia funcional e estrutural (Du et al., 2020; Trifonov et al., 2016). Estes eventos não são exclusivos do grupo supracitado, notavelmente eventos de duplicação genômica ocorreram dentro de Teleostei, sendo o Salmoniformes o grupo de maior destaque, na qual o evento ocorreu entre

50 - 100 ma em um processo de autopoloidização e com subsequente rediploidização (Dysin et al., 2022)

1.3 Importância da ordem Acipenseriformes

O táxon Actinopterygii é um dos grupos mais diversos de Vertebrata, compreendendo mais de 30 mil espécies válidas e representando mais da metade dos vertebrados descritos (Van Der Laan; Eschmeyer; Fricke, 2014). Esses organismos possuem ampla distribuição geográfica, ocupando praticamente todos os continentes e mares, e apresentam elevada diversidade morfofisiológica, além de grande importância comercial, biológica e para estudos evolutivos (Friedman, 2015).

Um dos grupos de destaque dentro de Actinopterygii é a família Acipenseridae, conhecida principalmente pelos esturjões, que têm grande importância econômica devido à produção de caviar, uma fonte de renda significativa para diversas comunidades há séculos (Bronzi; Rosenthal; Gessner, 2011). Entretanto, a intensa exploração desses peixes para consumo e produção de caviar resultou no declínio populacional de várias espécies, que agora constam na lista de espécies ameaçadas de extinção, agravado pela destruição de habitats devido às atividades antrópicas (Pikitch et al., 2005; Raposo et al., 2023).

Os Acipenseridae são encontrados exclusivamente no hemisfério norte e incluem os gêneros *Acipenser*, *Polyodon* e *Psephurus*. O gênero *Acipenser* abrange 25 espécies de esturjões, muitas com relevância econômica, enquanto *Polyodon spathula* e *Psephurus gladius* são espécies de grande porte, com desenvolvimento lento e hábitos diádromos ou potamódromos (Ludwig, 2008). Já o grupo externo é a família Polypteridae, única família que possui representantes atuais na base dos Actinopterygii, contendo 13 espécies do gênero *Polypterus*, incluindo a espécie *P. senegalus*, presente no estudo, e uma espécie do gênero *Erpetoichthys* (Near; Thacker, 2024)

Entre as espécies de interesse deste estudo, destacam-se: i) *Acipenser oxyrinchus*, encontrada no Atlântico Norte, principalmente na costa leste; ii) *Acipenser ruthenus*, restrita ao Mar Negro, majoritariamente de água doce, mas com populações estuarinas; iii) *Acipenser schrenckii*, do Nordeste da Ásia (Rio Amur), com hábitos de água doce e estuarinos; iv) *Acipenser sinensis*, presente no leste da Ásia (bacia do rio Yangtze, China), com comportamento anádromo; v) *Polyodon spathula*, espécie de água doce encontrada no interior

da América do Norte (Choudhury; Dick, 1998) e *Polypterus senegalus*, encontrado em grande parte da África equatorial (Lévêque, 1997).

Entender os elementos repetitivos no genoma, em especial os DNAs satélites, é fundamental para compreender a estrutura genômica, a evolução e auxilia na montagem de genomas com maior precisão (Shapiro & Sternberg, 2005; Tang et al., 2011; Tørresen et al., 2019). O presente estudo tem como objetivo compreender a história evolutiva dos satDNAs em um dos ramos basais dos Actinopterygii, grupo relevante e diversificado dentro dos vertebrados, visando compreender sua dinâmica evolutiva, associações com possíveis outros elementos e satDNAs conservados nos *taxa* estudados.

2. Objetivos

2.1. Objetivo Geral

Estudar a evolução de satDNAs em Actinopterygii basais, buscando compreender a diversidade, características, associações e dinâmica evolutiva.

2.2. Objetivos específicos

- Comparar todos os satelitomas caracterizados, visando encontrar DNAs satélites compartilhados no grupo;
- Quantificar a similaridade destes DNAs satélites e agrupá-los;
- Buscar os DNAs satélites altamente conservados em todas as espécies do estudo;
- Analisar os satDNAs conservados em *long-reads* ou genomas a nível cromossômico;

3. Material e métodos

3.1 Espécies utilizadas

As seguintes espécies da ordem Acipenseriformes foram analisadas no presente trabalho: *Polyodon spathula*, *Acipenser oxyrinchus*, *Acipenser schrenckii*, *Acipenser sinensis* e *Acipenser ruthenus*, e um representante da família Polypteridae foi utilizado como grupo externo, a espécie *Polypterus senegalus*. Todas estas bibliotecas estão de acesso público na plataforma NCBI, e o número de acesso para cada biblioteca, junto com demais dados referente às bibliotecas e aos genomas dos organismos estão sumarizados na Tabela 1.

Tabela 1. Características genômicas das espécies estudadas, com o tamanho estimado do genoma, e o tamanho total da biblioteca, em gibabases, e o respectivo código de acesso na base de dados “Sequence Read Archive – SRA”

Espécie	Tamanho estimado do genoma (gb)	Tamanho total da biblioteca (gb)	Cobertura	Código de acesso
<i>Polypterus senegalus</i>	Sem dados	11,2	-	SRR11149362
<i>Polyodon spathula</i>	1,55	4,70	3,03×	SRR11149443
<i>Acipenser oxyrinchus</i>	2,14	1,09	0,5×	SRR11679494
<i>Acipenser schrenckii</i>	Sem dados	16,85	-	SRR11183103
<i>Acipenser sinensis</i>	Sem dados	0,87	-	SRR6327824
<i>Acipenser ruthenus</i>	1,81	6,30	3,48×	SRR8371834

3.2 Caracterização dos satelitomas

A caracterização de todos os satelitomas das espécies começa com o download das bibliotecas *paired-end* do SRA e um pré-processamento, iniciando com a checagem de sua qualidade através do FastQC (Andrews, 2010) utilizando um servidor de alta capacidade, aos cuidados do Prof. Dr. Claudio de Oliveira, do Instituto de Biociências da UNESP – Botucatu, e depois as bibliotecas foram avaliadas quanto à qualidade e tamanho dos *reads* utilizando o software Trimmomatic (Bolger; Lohse; Usadel, 2014) visando remover *reads* de baixa

qualidade e cortar aqueles que sobraram para 101 pares de base, o padrão para as análises subsequentes.

Para isolar propriamente o satelitoma das espécies, foi utilizado o *software RepeatExplorer2* (Novák; Neumann; Macas, 2020), na qual foi realizado uma subamostragem de 500.000 *reads* para cada biblioteca, totalizando 1.000.000 *reads*, e na primeira iteração foi utilizado o *script* *rexp_prepare_normaltag* (Ruiz-Ruano et al., 2016) para preparar essa biblioteca para ser enviada ao servidor Galaxy, que foi utilizado até a finalização das iterações no *RepeatExplorer2* (The Galaxy Community, 2024). No servidor, o *RepeatExplorer2* foi rodado com as configurações padrões, a única mudança foi a alocação de memória RAM para “*Extra-Long*”, com as sequências obtidas, e para as iterações subsequentes, essas sequências foram utilizadas para filtrar a biblioteca original de modo a permitir que o *RepeatExplorer2* conseguisse obter novos satDNAs utilizando o *software* DeconSeq (Schmieder; Edwards, 2011) e o *script* *rexp_prepare_deconseq.py* (Ruiz-Ruano et al., 2016). A biblioteca resultante, após a remoção dos *reads* pertencentes a sequências de iterações anteriores, foi enviada para o Galaxy novamente, o *RepeatExplorer2* foi rodado nas mesmas configurações, e o ciclo continuou até que não fosse possível isolar mais nenhum satDNA.

Com este catálogo bruto foram removidas as famílias multigênicas, sequências altamente conservadas entre os diversos *taxa* de Vertebrata, utilizando o *software* Geneious 7.1 usando sua função de anotação, que busca similaridade entre uma base de dados (sequências de multigênicas) e sequências alvo (satDNAs), também foi realizado a busca por similaridade entre as sequências utilizando o RepeatMasker (Smit, A. F. A.; Hubley, R.; Green, P., 2015) e a opção “*de novo assembly*” do Geneious, de forma a retirar-se satélites iguais (> 95% de similaridade), de encontrar as variantes de uma mesma família (> 80% de similaridade e < 95%), sendo que estes elementos são contabilizados na abundância, mas removidos do catálogo final, e agrupando aqueles que possuem uma relação de similaridade > 50% em Super Famílias (Ruiz-Ruano et al., 2016).

Para o cálculo de abundância e divergência dos satelitomas foi utilizado o *software* RepeatMasker (Smit, A. F. A.; Hubley, R.; Green, P., 2015) utilizando 5.000.000 de *reads*, totalizando 10.000.000. Quando não foi possível utilizou-se o máximo de *reads* disponível para a espécie. De forma a padronizarmos a nomenclatura dos satDNA foi utilizada a convenção proposta por Ruiz-Ruano e colaboradores (2016) que consiste em nomear os satélites utilizando a primeira letra do gênero e as duas primeiras do epíteto específico, a sigla Sat, seguido pelo

número do satélite, em ordem decrescente de abundância, e pelo tamanho da unidade de repetição, como exemplo, o primeiro satélite da espécie *Acipenser oxyrinchus* é o AoxSat01-4200. A tabela 2 contém informações gerais sobre todos os satelitomas caracterizados para as espécies, e as tabelas completas estão no material suplementar (Tabelas S1 – S6).

3.2 Satelitômica comparativa e BLAST

Visando melhor compreensão da coleção de satDNAs compartilhado entre as espécies foi realizado um estudo comparativo usando o catálogo completo de todas as espécies do estudo, em uma primeira etapa a opção “*de novo assembly*” do Geneious foi utilizada para filtrar potenciais similaridades, essas sequencias foram alinhadas para identificar as similaridades e, a posteriori, o *RepeatMasker* para comparar os catálogos de satDNAs. Para isso, todos os satelitomas foram agrupados em um arquivo fasta e foi realizado uma análise para encontrar satDNAs potencialmente similares, na qual todos os satDNAs foram mantidos, exceto aqueles que apresentavam mais de 80% de similaridade, na qual apenas um representante foi mantido no fasta final, conforme indicado nas tabelas S7 e 3. Após isso, este fasta foi utilizado no *RepeatMasker* contra as bibliotecas de cada estudo, e a posteriori os resultados de abundância e divergência dos satDNAs encontrados foram plotados em gráficos do tipo *RepeatLandscape*, utilizando um script previamente estabelecido por Ruiz-Ruano et al. (2016) denominado `satminer_quant.py` (https://github.com/fjrui.ruano/ngs-protocols/blob/master/satminer_quant.py) gerando assim gráficos de proporção genômica e a Divergência-2-Kimura para todos os catálogos (Figuras 1 - 6).

Para a análise via BLAST (Sayers et al., 2025) foi utilizado a versão *web* com os satDNAs da tabela 3, coluna 1, e os seguintes parâmetros para a busca: *standard database*; “*highly similar sequences (megablast)*” e as demais opções, no BLAST, foram mantidos no padrão; de forma a filtrar os resultados, e buscar apenas aqueles mais específicos e com valor biológico, os resultados foram filtrados para: *Query coverage* $\geq 95\%$ e *e-value* o menor possível, no mínimo 1^{-09} . Estes resultados nos levaram a Scaffolds de sequenciamento para as respectivas espécies, que foram utilizados para analisar a região que o satDNA foi dado como resultado do BLAST e todos os resultados obtidos foram sumarizados na tabela 4. De forma a comparar um dos resultados obtidos as seguintes sequências do DNA satélite *HindIII* foram utilizadas para fazer um alinhamento: *Acipenser transmontanus*: AJ286603.1, *A. brevirostrum*: AJ555334.1, *A. stellatus*: AJ555347.1.

Para analisar os *Scaffolds* e a possível presença de satDNAs foi utilizado a função "Genome data viewer" do NCBI, na qual é possível visualizar os *Scaffolds* que contém a anotação do gene previsto, extrair essa região e visualizá-la, a possível presença do respectivo satDNA é observada utilizando o Geneious, com sua função de anotação. Com isso, obteve-se as seguintes regiões: *long non-coding* RNA (XR_009323890.1) do AoxSat01-4200, as regiões dos mRNA para os satélites AoxSat14-52, AscSat25-57, AscSat27-43 e AscSat28-80 (respectivamente: XM_059020450.1, XM_059018971.1, XM_059021770.1, XM_059020903.1).

Para identificar a região que o DNA satélite pudesse estar foi observado o *DotPlot* do DNA satélite contra a sequência supracitada, e assim regiões que possuem a repetição foram extraídas e alinhadas, usando a configuração padrão do algoritmo de alinhamento *Global Alignment with free end gaps* do Geneious 7.1, além de utilizar a função de anotação do Geneious para poder marcar esses DNAs satélites nos genes previstos. Tanto esses genes como os alinhamentos foram salvos e montados como as figuras 3, 4 e S8 com o Photoshop CS6.

4. Resultados

4.1 Caracterização dos satelitomas

4.1.1 *Polyodon spathula*

Foram realizadas 6 iterações no *RepeatExplorer2* (RE2) para a espécie, totalizando 69 possíveis satDNAs caracterizados. Após isso, uma análise em busca de famílias multigênicas, elementos que possam vir a ser isolados como satDNAs, foi realizada e nenhuma família foi encontrada. Na análise de similaridade foram retirados 4 satélites, totalizando 65 satDNAs para a espécie, com a presença de 4 superfamílias contendo 27 satDNAs nestas (Tabela 2, Tabela S1), e para a análise de abundância e divergência dessas sequências foi realizado um subsample de 4.500.000 de reads para cada biblioteca (totalizando 9.000.000) e obtivemos os dados que geraram a Tabela final do satelitoma (Tabela S1) bem como o *RepeatLandscape*, que apresenta a abundância dos satDNAs e a divergência por nível de substituição de Kimura (Figura S1).

Tabela 2. Características gerais dos satelitomas caracterizados neste estudo. **RUL:** *Repeat unit length*.

Espécie	satDNAs	Super Famílias	Mediana da RUL	Média da abundância (%)	Média da divergência (%)	Média do conteúdo A+T (%)
<i>Polyodon spathula</i>	65	4 (27 satDNAs)	178	0.0004176	6.53	55.31
<i>Acipenser oxyrinchus</i>	34	2 (4 satDNAs)	51.5	0.001258	7.84	52.38
<i>Acipenser schrenckii</i>	42	4 (10 satDNAs)	153	0.001027	7.29	56.02
<i>Acipenser sinensis</i>	53	5 (11 satDNAs)	39	0.000448	7.09	47.79
<i>Acipenser ruthenus</i>	30	1 (2 satDNAs)	59	0.000619	6.16	52.08
<i>Polypterus senegalus</i>	09	-	485	0.000763	5.65	57.61

4.1.2 *Acipenser oxyrinchus*

Após três iterações no RE2 foram encontrados 36 satDNAs para a espécie, na busca por famílias multigênicas foram retiradas duas sequências, e nas análises de similaridade não foi encontrado nenhuma variante ou satDNA igual, totalizando 34 satDNAs para a espécie, com a presença de 2 SF com 4 satDNA nelas (Tabela 2, Tabela S2). A análise de abundância e divergência foi realizada utilizando *subsamples* de 4.000.000, totalizando 8.000.000 *reads* total, com tanto a tabela final do satelitoma sendo gerada (Tabela S2) como o *RepeatLandscape* (Figura S2).

4.1.3 *Acipenser schrenckii*

Nesta espécie foram realizadas sete iterações no RE2 para caracterizar-se 44 possíveis satDNAs, com a análise de famílias multigênicas foram retiradas duas sequências, totalizando 42 sequencias caracterizados. A análise de similaridade não apresentou nenhum elemento como variantes tampouco mesmo satDNA, apenas quatro SF com 10 satDNAs nelas (Tabela 2, Tabela S3). Na análise de abundância e divergência 10.000.000 *reads* foram utilizados, permitindo a construção da tabela para a espécie (Tabela S3) e a geração do *RepeatLandscape* (Figura S3).

4.1.4 *Acipenser sinensis*

Foram caracterizados 54 satDNAs nas seis iterações do RE2, sem a presença de famílias multigênicas para esta espécie. Com um caso de mesma sequência sendo caracterizada mais de uma vez, a espécie apresenta um satelitoma contendo 53 satDNAs, com 5 SF que possuem 11 satDNAs distribuídos nelas (Tabela 2). A análise de abundância e divergência, realizadas com um *subsample* de 10.000.000 de *reads*, permitiu a finalização da tabela desta espécie (Tabela S4) e a construção do *RepeatLandscape* dos satDNAs (Figura S4).

4.1.5 *Acipenser ruthenus*

Com três iterações no RE2 foram isolados 32 possíveis satDNAs, que, na análise de presença de famílias multigênicas, foi encontrado um elemento, além de uma variante de uma mesma família na análise de similaridade realizada, logo, o catálogo desta espécie compreende 30 satDNAs caracterizados, com uma SF com dois satDNAs (Tabela 2). Um *subsample* de 10.000.000 *reads* foi necessário para os cálculos de abundância e divergência, que gerou os dados presentes na Tabela do satelitoma desta espécie (Tabela S5) bem como no *RepeatLandscape* (Figura S5).

4.1.6 *Polypterus senegalus*

Após três iterações no RE2 foram isolados 11 possíveis satDNAs nesta espécie que, após a análise da presença de multigênicas, na qual uma sequência foi retirada, e na análise de similaridade, na qual mais uma sequência foi retirada, contém 9 satDNAs caracterizados, sem a presença de Super Famílias (Tabela S6). Utilizou-se um subsample de 10.000.000 reads para os cálculos de abundância e divergência, com os dados obtidos sendo encontrados na Tabela S6 e na Figura S6.

4.2 Satelitômica comparativa na ordem Acipenseriformes

A análise comparativa foi feita utilizando todos os satelitomas descritos para as espécies deste estudo e relevou a presença de 106 famílias de satDNA que apresentam similaridade ($\geq 50\%$), resultando em 28 famílias conservadas de satDNA (Tabela S8). A maioria dos satDNAs presente nas famílias conservadas foram caracterizados em membros do gênero *Acipenser* e compartilham similaridades entre si, com apenas 18 satDNAs deste grupo das famílias conservadas que foram caracterizados em *Polyodon spathula* e não foram encontrados satDNAs similares em *Polypterus senegalus*. Ao observar os satDNAs conservados dentro do grupo foi possível constatar que existe um predomínio de satDNAs curtos (< 100 pb; 16 no total) em relação aos longos (> 100 pb; 8; tabela 3), com todos eles possuindo mais que 50% de

composição A+T, com a maioria sendo satDNAs com uma abundância média/baixa nos genomas na qual foram caracterizados e com uma baixa taxa de divergência.

Tabela 3. satDNAs conservados dentro da ordem Acipenseriformes.

satDNA I	satDNA II	% de Similaridade
AoxSat14-52	AscSat20-52	100
AscSat28-80	AoxSat16-80	100
AruSat08-31	AscSat38-31	100
AscSat25-57	PspSat23-57	98.246
AscSat26-28	AsiSat25-28	100
AscSat27-43	PspSat46-43	100
AscSat41-187	AoxSat11-187	99.465
AscSat07-271	AsiSat01-271	99.262
AscSat02-171	AsiSat21-171	98.83
AoxSat01-4200	AscSat01-4200	98.571
AruSat30-36	PspSat26-36	97.222
AsiSat29-40	PspSat63-40	95

Para compreender a possível presença destes satDNAs conservados nas outras espécies do estudo foi realizado uma análise comparativa com todos os satDNAs, exceto aqueles que apresentavam similaridade superior à 80%, contra todas as bibliotecas utilizadas, gerando *landscapes* com a cobertura dos satDNA nas bibliotecas. De uma maneira geral, observando os gráficos para as espécies do gênero *Acipenser* foi possível observar que as bibliotecas de satDNAs do gênero se encontram representadas nos genomas do gênero, com altas quantidades de sequências com baixa taxa de divergência (0 – 5% em níveis de substituição de Kimura), enquanto *Polyodon spathula*, o grupo externo do gênero, também foi encontrado, mas com taxas maiores de divergência nas espécies (0 - >5% nas espécies de *Acipenser*).

Observando apenas o gráfico de *P. spathula* foi possível observar que, além de seus próprios satDNAs, as sequências encontradas no gênero também estão presentes nesta espécie, mas com divergências maiores que as encontradas no gênero, conforme previsto pela hipótese *Library*, indicando um processo de homogeneização e fixação ocorrendo para certas sequências, que deram origem a satDNAs com baixas taxas de divergências nas espécies do gênero *Acipenser*, e que este conjunto de satDNAs possui uma origem anterior à separação de *P. spathula* e de *Acipenser*. Em contrapartida, o gráfico de *Polypterus senegalus* indica presença quase exclusiva de sequências apenas desta espécie, com baixíssima representação de sequências de outras espécies que, provavelmente, são artefatos da análise, como alinhamento

de regiões pequenas e com alguma similaridade com a referida biblioteca, visto que em nenhuma outra análise de similaridade obteve resultado que demonstrasse a presença de satélites compartilhado, o que indica que não houve origem das sequências compartilhadas entre Acipenseriformes nesta espécie.

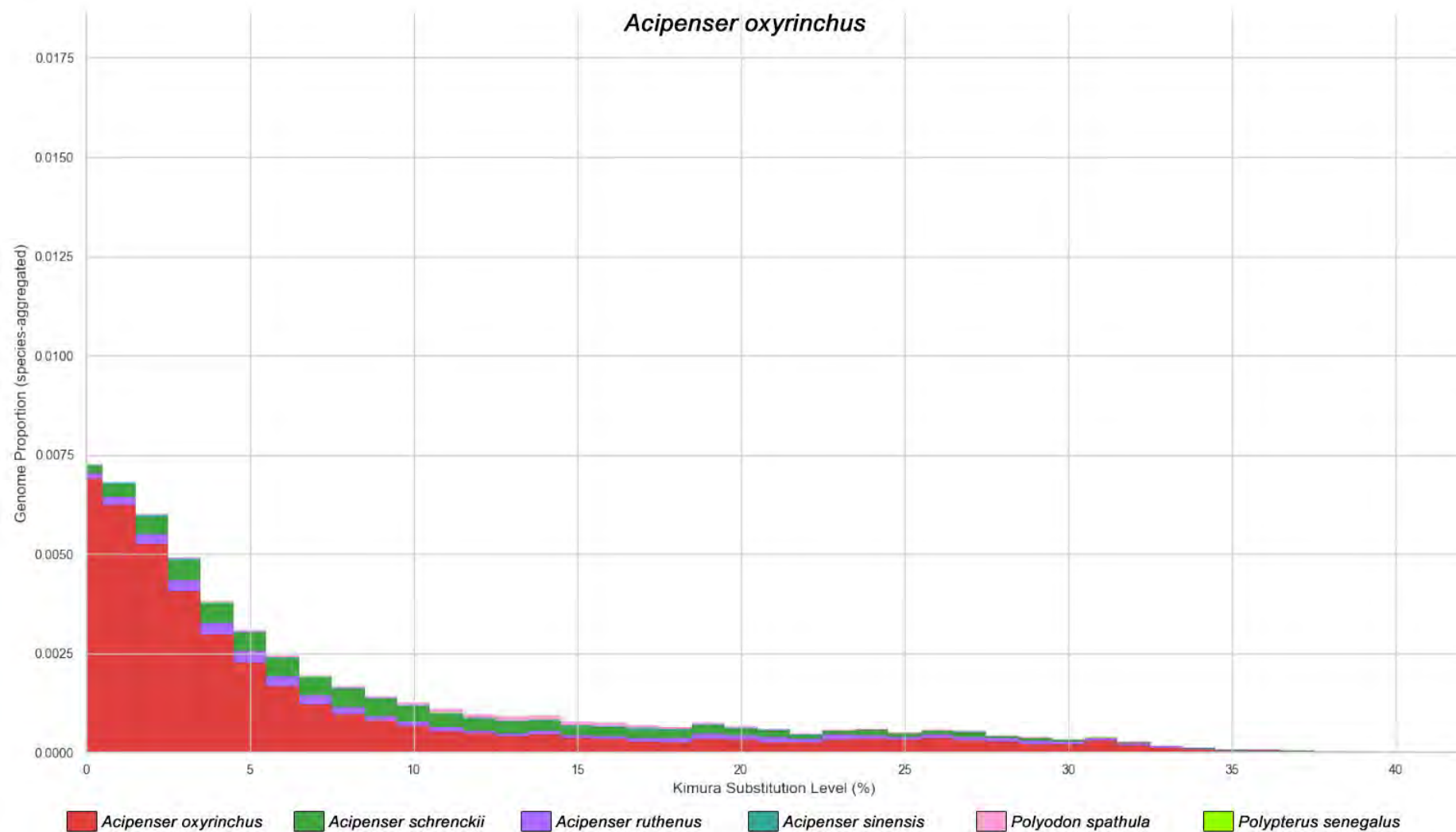


Figura 1. *RepeatLandscape* de grande parte dos satDNAs do estudo contra o genoma de *Acipenser oxyrinchus*. Cada cor representa o catálogo parcial de satDNAs da espécie.

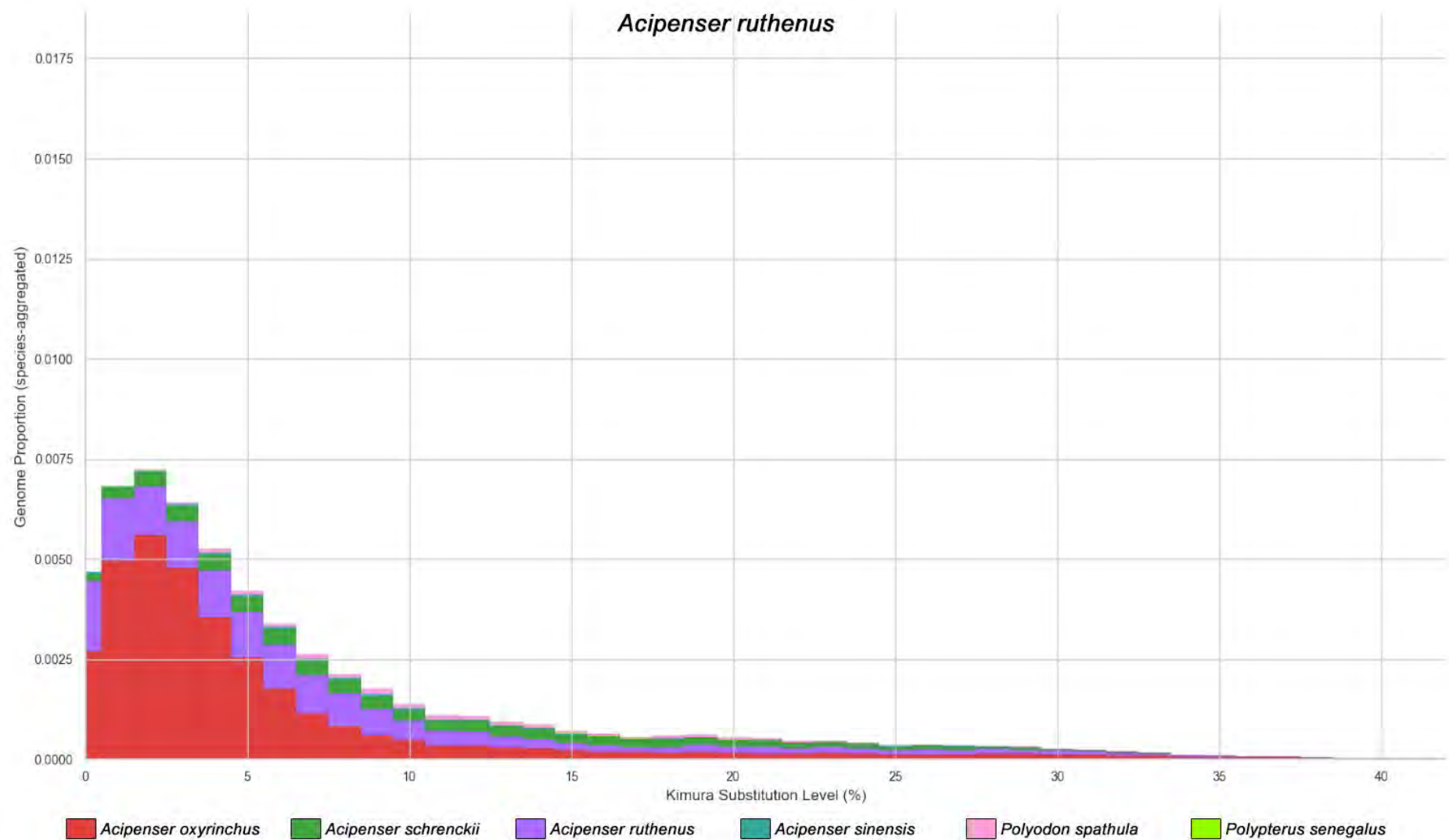


Figura 2. RepeatLandscape de grande parte dos satDNAs do estudo contra o genoma de *Acipenser ruthenus*. Cada cor representa o catálogo parcial de satDNAs da espécie.

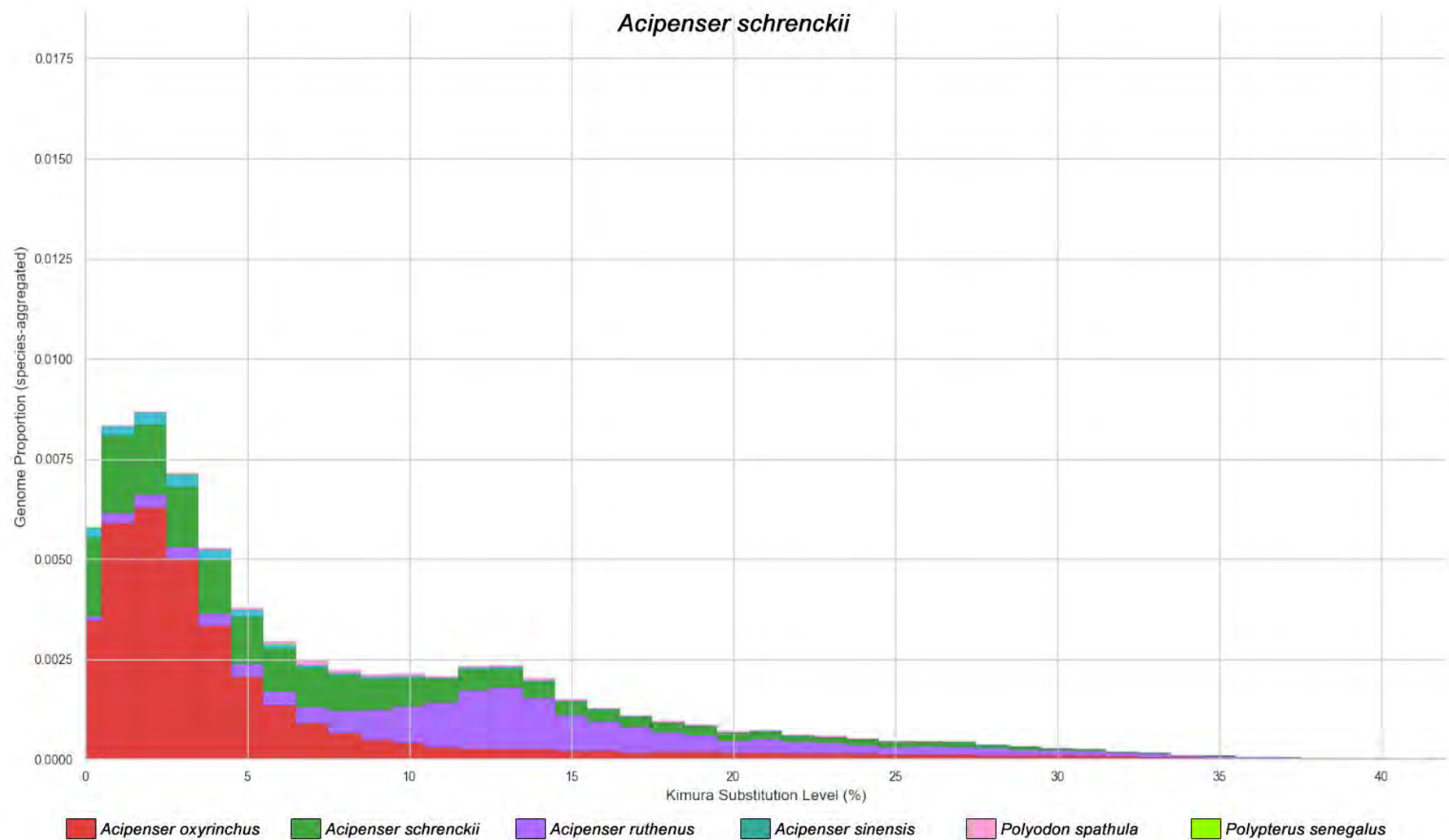


Figura 3. *RepeatLandscape* de grande parte dos satDNAs do estudo contra o genoma de *Acipenser schrenckii*. Cada cor representa o catálogo parcial de satDNAs da espécie.

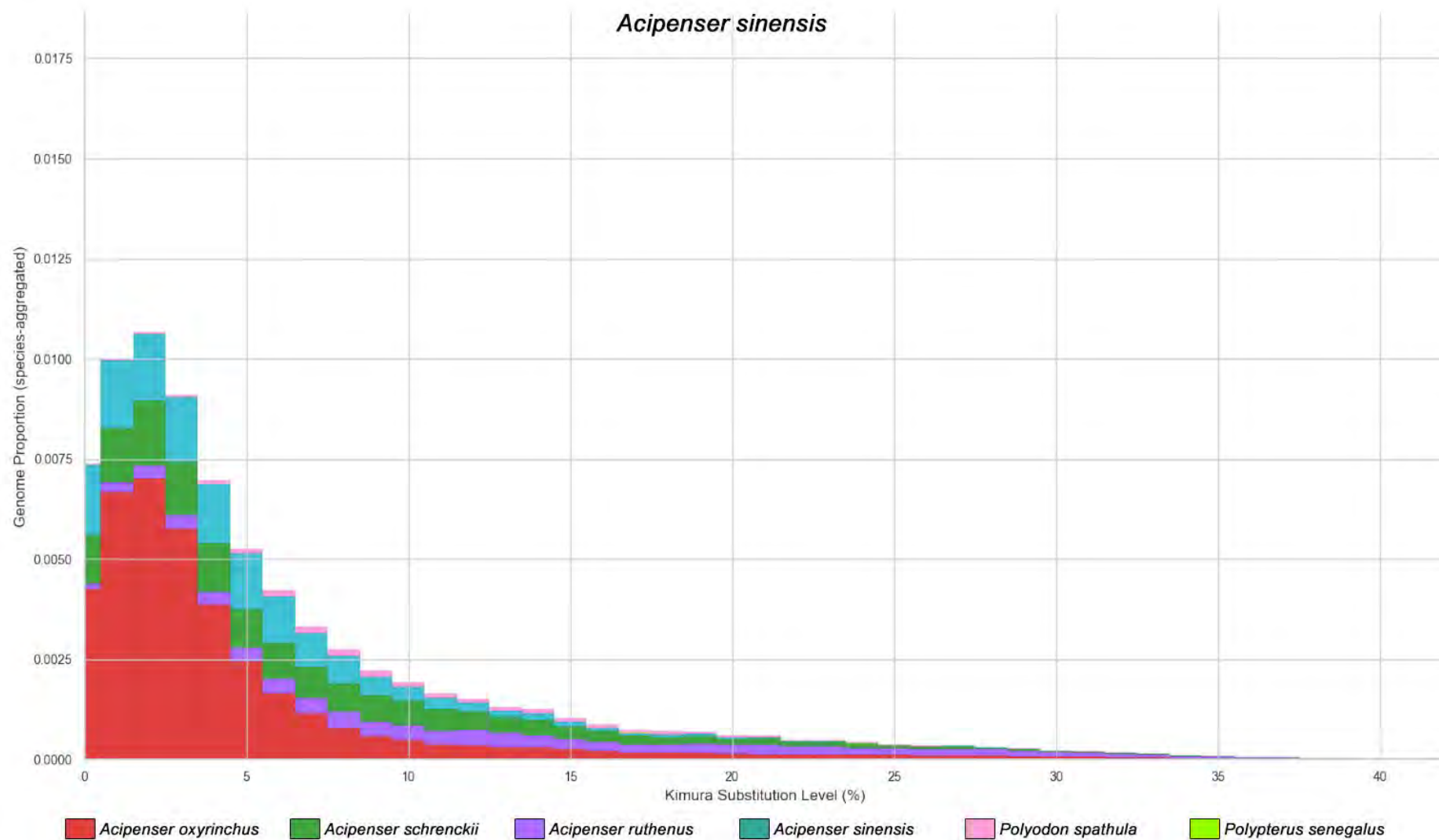


Figura 4. *RepeatLandscape* de grande parte dos satDNAs do estudo contra o genoma de *Acipenser sinensis*. Cada cor representa o catálogo parcial de satDNAs da espécie.

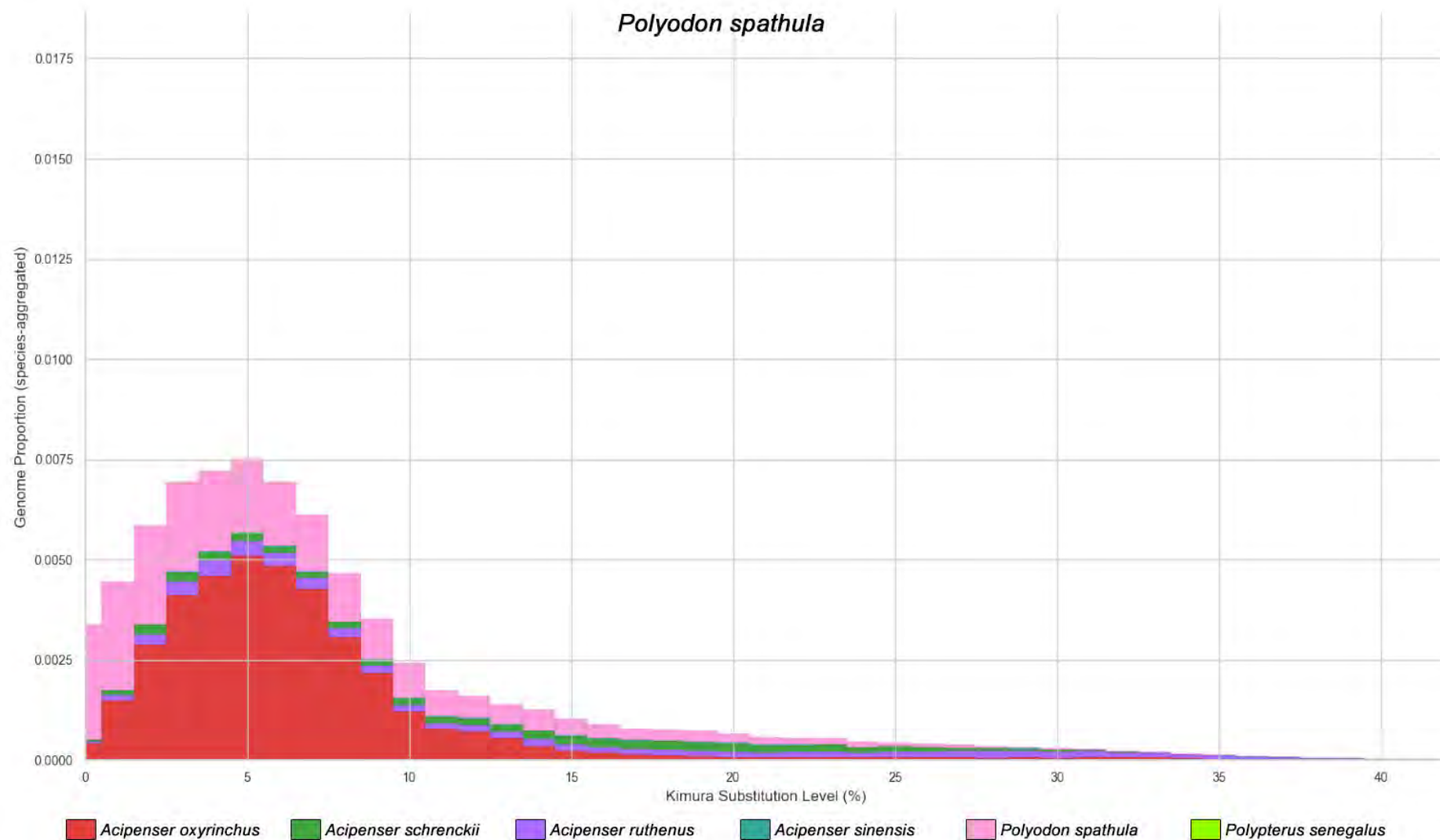


Figura 4. *RepeatLandscape* de grande parte dos satDNAs do estudo contra o genoma de *Polyodon spathula*. Cada cor representa o catálogo parcial de satDNAs da espécie.

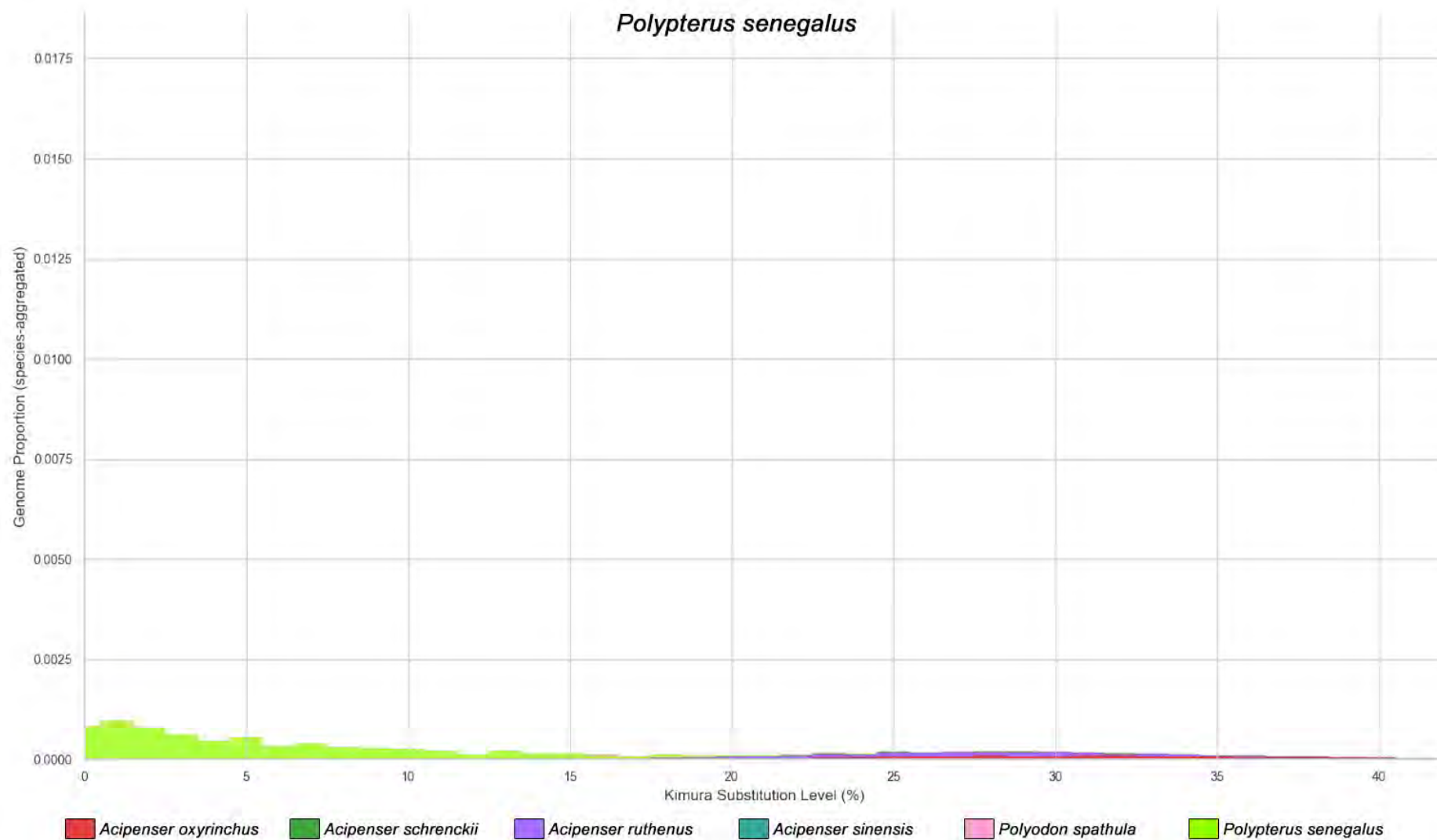


Figura 5. *RepeatLandscape* de grande parte dos satDNAs do estudo contra o genoma de *Polypterus senegalus*. Cada cor representa o catálogo parcial de satDNAs da espécie.

4.3 Resultados do BLAST

A comparação das sequências altamente conservadas com a base de dados do NCBI, via BLAST, trouxe como resultado seis DNAs satélites com algum tipo de resultado considerável, com o AoxSat01-4200 sendo extremamente similar com um possível lncRNA, os DNAs Satélites AoxSat14-52, AscSat25-57, AscSat27-43 e AscSat28-80 apresentando similaridade com diversas classes de possíveis mRNA e com o AscSat02-171 apresentando similaridade com o DNA satélite *HindIII*.

Tabela 5. Resultados sumarizados do BLAST. Como em alguns casos a quantidade de dados era alta, apenas um resultado foi apresentado na coluna “**Resultados do BLAST**”, com a sequência de maior valor sendo descrita, com o **código de acesso** para determinada sequência e na coluna **organismos** outras espécies que tiveram *hit* que condizem com os parâmetros estabelecidos para a análise.

satDNA I	Resultado do BLAST	Organismos:	Código de acesso
AoxSat01-4200	PREDICTED: <i>Acipenser ruthenus</i> uncharacterized LOC131730125 (LOC131730125), ncRNA	<i>Acipenser ruthenus</i> - 1 hit.	XR_009323890.1
AoxSat14-52	Diversos mRNA, como exemplo: PREDICTED: <i>Acipenser ruthenus</i> guanine nucleotide binding protein (G protein), gamma 3 (gng3), transcript variant X4, mRNA	<i>Acipenser ruthenus</i> - 91 hits; <i>Polyodon spathula</i> - 9 hits.	XM_059020450.1
AruSat08-31	Não houve correspondência	-	-
AruSat30-36	Não houve correspondência	-	-
AscSat02-171	Correspondência com o DNA satélite <i>HindIII</i> , como: <i>Acipenser transmontanus HindIII</i> satellite DNA family, clone pTR1	<i>Acipenser transmontanus</i> - 3 hits; <i>A. brevirostrum</i> - 1 hit; <i>A. stellatus</i> - 1 hit.	AJ286603.1
AscSat07-271	Não houve correspondência	-	-
AscSat25-57	Diversos mRNA, como exemplo: PREDICTED: <i>Acipenser ruthenus</i> uncharacterized LOC117968050 (LOC117968050), transcript variant X1, mRNA	<i>Acipenser ruthenus</i> - 24 hits; <i>Acipenser brevirostrum</i> - 1 hit; <i>Polyodon spathula</i> - 5 hits.	XM_059018971.1
AscSat26-28	Não houve correspondência	-	-
AscSat27-43	Diversos mRNA, como exemplo: PREDICTED: <i>Acipenser ruthenus</i> CD48 antigen-like (LOC117409995), transcript variant X1, mRNA	<i>Acipenser ruthenus</i> - 12 hits; <i>Polyodon spathula</i> - 2 hits.	XM_059021770.1
AscSat28-80	Diversos mRNA, como exemplo: PREDICTED: <i>Acipenser ruthenus</i> cathepsin D-like (LOC131731654), mRNA	<i>Acipenser ruthenus</i> - 84 hits; <i>A. baerii</i> - 1 hit; <i>Polyodon spathula</i> - 12 hits.	XM_059020903.1
AscSat41-187	Não houve correspondência	-	-
AsiSat29-40	Não houve correspondência	-	-

Devido à similaridade encontrada AscSat02-171 com o DNA satélite *HindIII* foi realizado um alinhamento deste satélite (Figura 2) com três sequências isoladas em outras espécies de esturjões para quantificar a similaridade apresentada, na qual foi observado que existe um alto grau de similaridade, com *Acipenser transmontanus* apresentando o maior valor

de similaridade, com 97.661%, seguido de *Acipenser brevirostrum*, com 91.813% e com *Acipenser stellatus* com 89.474%.

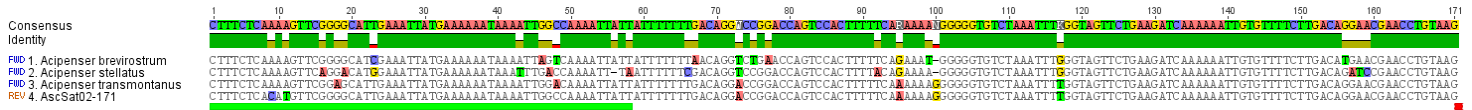


Figura 6. Alinhamento de AscSat02-171 com sequências isoladas de *Hind*III de outras espécies do gênero *Acipenser*.

Quatro regiões possivelmente transcritas foram identificadas nos *scaffolds* de *Acipenser ruthenus*, uma região de *long non-coding RNA* (lncRNA) que teve similaridade com o DNA satélite AoxSat01-4200, quatro regiões, em *scaffolds* distintos, com mRNAs diferentes, como do transcrito *guanine nucleotide binding protein (G protein)*, *gamma 3 (gng3)*, *transcript variant X4*, mRNA para o AoxSat14-52, um mRNA não caracterizado para AscSat25-57, o transcrito *CD48 antigen-like transcript variant XI* mRNA para AscSat27-43 e o transcrito *cathepsin D-like* para AscSat28-80, sendo todos esses resultados previstos pelo algoritmo de predição do NCBI, não sendo dados efetivamente observados por experimento.

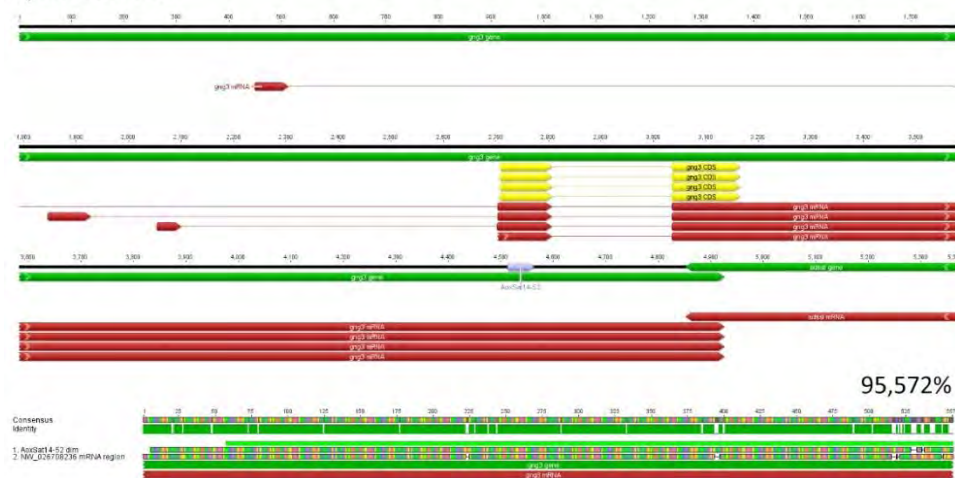
A região do AoxSat01-4200 é uma região na qual o DNA satélite está praticamente inteiro inserido em lncRNA, com 4,179 pares de base presentes no final da região, (Figura 3a) com ~90% de similaridade com essa região; o AoxSat14-52 está presente apenas em uma região do transcrito, com 95,572% de similaridade com a região que ele está presente (Figura 3b). O AscSat25-57 está presente em uma longa região do transcrito (1,771 pares de base), onde o mesmo está repetido em tandem, com 96,842% de similaridade (Figura 3c).

O AscSat27-43 está presente em uma região do transcrito (Figura 4b-i) com 86,068% e, além disso, outro DNA satélite foi encontrado, o AscSat25-57 (figura 4b-ii), previamente descrito em outro transcrito, presente em tandem nessa região de ~3.500 pares de base, mas já apresentando uma divergência em relação à sequência do transcrito, com o alinhamento possuindo 74,246% de similaridade com o DNA satélite; por fim, o AscSat28-80 está presente em uma região de ~600 pares de base, também em tandem, com uma similaridade de 96,696% (figura 4a); todas as regiões encontradas na análise do *dotplot* foram extraídas e alinhadas, com os alinhamentos presentes na Figura S7.

a) AoxSat01-4200



b) AoxSat14-52



c) AscSat25-57

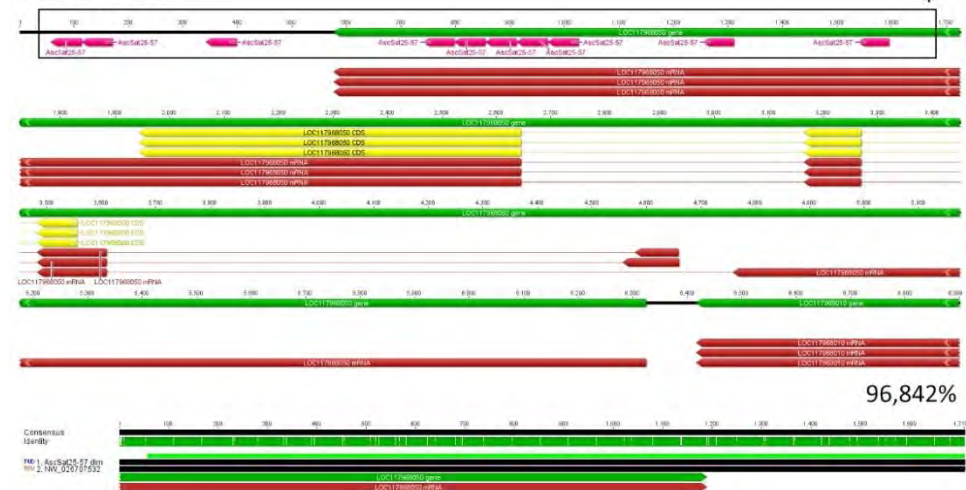
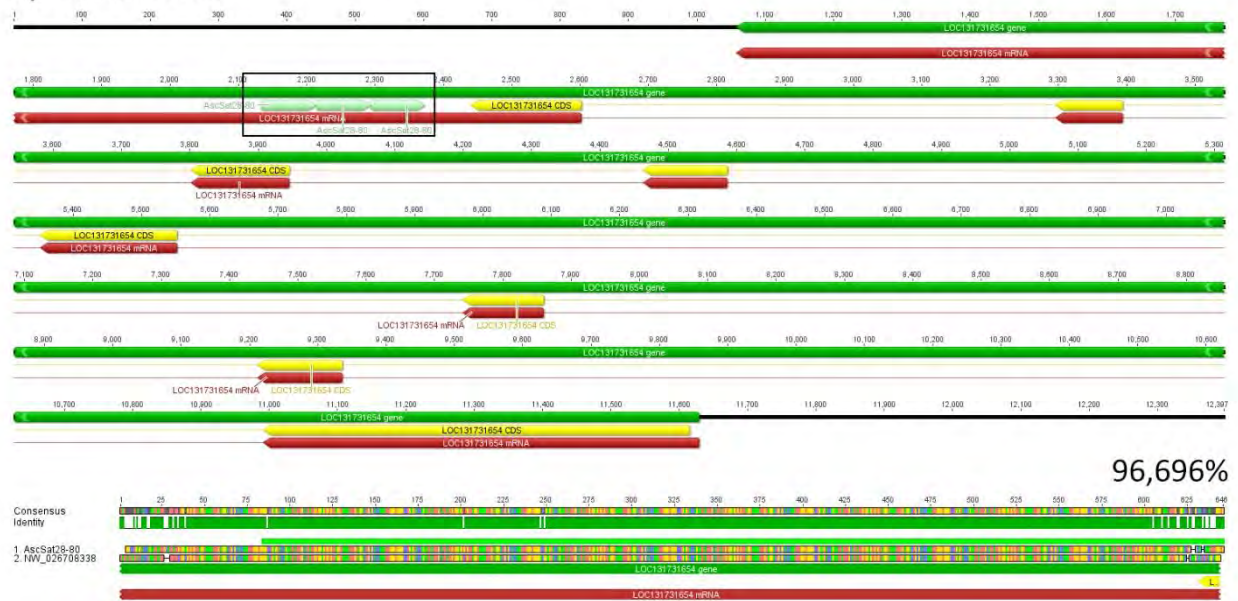


Figura 7. As regiões extraídas dos *scaffolds* e o alinhamento com o satDNA demonstrando a presença dele. a) AoxSat01-4200 no lncRNA; b) AoxSat14-52 com o mRNA; c) AscSat25-57, com o asterisco indicando a região que ele se encontra em tandem no possível transcrito.

a) AscSat28-80



b) AscSat27-43



Figura 8. As regiões extraídas dos *scaffolds* e o alinhamento com o satDNA demonstrando a presença dele. a) AscSat28-80 com a região do DNA satélite destacada, mostrando que ele está em tandem neste possível transcrito. b) AscSat27-43 na região do possível transcrito; é possível observar que existem dois DNAs Satélites neste transcrito, em *i* observa-se o AscSat27-43 numa região final do transcrito, e em *ii* observa-se uma longa região, em tandem, do DNA Satélite AscSat25-57.

5. Discussão

No presente contexto da Biologia, a genômica é uma área que vem se expandindo e se tornando cada vez mais importante, seja no âmbito médico, de pesquisa e desenvolvimento, ambiental ou, no contexto do presente estudo, evolutivo, e essa importância é multifacetada, onde os diversos elementos do genoma são importantes para se ter uma melhor compreensão do genoma eucarioto e sua evolução, estruturação e funcionamento.

Os DNAs satélites se tornam elementos importantes, pois sua descrição e compreensão auxiliam tanto em processos de montagem de genoma, como permite compreender as características particulares deste elemento, que por muito tempo levantou mais questões do que respostas, mas com o advento da genômica, foram sendo explorados e mais bem compreendidos. Logo, compreender a evolução de satDNAs em Acipenseriformes, uma das ordens basais dentro de Actinopterygii, é construir um arcabouço que vai permitir o entendimento da dinâmica evolutiva de satDNAs em um dos *taxa* mais ricos, abundantes e com uma ampla diversidade tanto de espécies como de características genômicas de Vertebrata, na qual já existe estudos comparativos em Characoidei, descrevendo inclusive a expressão do CharSat01, um satDNA conservado (Dos Santos et al., 2021).

Neste estudo foi realizado a descrição de seis novos satelitomas em Actinopterygii, 4 satelitomas para o gênero *Acipenser*, nas espécies *A. oxyrinchus*, *A. schrenckii*, *A. sinensis*, *A. ruthenus*, e um satelitoma para o gênero *Polyodon*, em *P. spathula*, grupo irmão do gênero *Acipenser*, e de *Polypterus senegalus*, grupo externo de Acipenseriformes. Como já descrito em outros satelitomas dentro de Teleostei, a abundância dos satDNAs no genoma é baixa, não sendo maior que 0.001% (Goes et al., 2023; Kretschmer et al., 2022; Utsunomia et al., 2017), diferente do que acontece em outros *taxa*, como no gênero *Drosophila*, na qual o satelitoma de certas espécies pode ultrapassar 25% (De Lima; Ruiz-Ruano, 2022).

A satelitômica comparativa revelou características interessantes para o grupo, o primeiro caso é a alta quantidade de satDNAs compartilhados, principalmente entre os membros do gênero *Acipenser*, foram encontrados 106 satDNAs compartilhados entre os *taxa*, com grande parte deles sendo compartilhados dentro do gênero *Acipenser*, o que corrobora a hipótese *Library* (Fry; Salser, 1977) que, em um dos seus pressupostos, diz que cada espécie contém uma biblioteca de satDNAs que podem ser compartilhados, mas vão apresentar padrões de expansão e contração diferentes nas espécies, o que é observado no presente estudo, com

satDNAs compartilhados, mas com valores de abundância que diferem entre as espécies, como também é observado em outros satelitomas descritos para peixes (Goes et al., 2022, 2023; Stornioli et al., 2021), como também em outros grupos, como em Coleoptera (Mora et al., 2024), em aves (Setti et al., 2024) e em Crocodylia (Sales-Oliveira & dos Santos et al., 2024), sendo este último grupo particular por apresentar algumas características similares com a ordem Acipenseriformes, tempo de divergência parecido e uma lenta evolução molecular. A conservação destes elementos ao longo de milhões de anos pode ocorrer por uma possível função estrutural ou funcional (Šatović-Vukšić; Plohl, 2023) ou devido à própria evolução molecular lenta neste grupo de organismos (Robles et al., 2004), sendo necessário uma investigação mais aprofundada nestes aspectos.

As características gerais dos satDNAs compartilhados demonstra que existe um predomínio de satDNAs curtos em relação aos longos, possivelmente indicando que alguns satDNAs curtos se originaram *de novo* no ancestral em comum, visto que existe uma hipótese de que os curtos possam se originar *de novo* em uma dada espécie (Camacho et al., 2022) e que, com o passar do tempo, aumente sua complexidade e tamanho do monômero (Stephan; Cho, 1994). Todos os satDNAs conservados apresentaram baixa divergência, o que é explicado pelo fato de que mesmo possuindo uma evolução molecular e homogeneização lenta este evento vêm acontecendo há pelo menos 150 ma, para aqueles satDNAs que não se originaram *de novo* no gênero *Acipenser*, e há 100 ma (Shen et al., 2020) para aqueles que se originaram *de novo*, tempo suficiente para que essas sequências possam ser homogeneizadas (De La Herrán et al., 2001). Conforme já observado em trabalhos de outro grupo de peixes, os Teleostei, existe um viés de a média de conteúdo A+T neste grupo ser maior que 50%, como exemplo existem satelitomas com 55% (Silva et al., 2017) e até mesmo 65% (Serrano-Freitas et al., 2020) de conteúdo A+T, na qual é possível observar este viés na base de Actinopterygii, com a única exceção sendo *Acipenser sinensis* que possui o conteúdo A+T de aproximadamente 48%.

A análise de similaridade permitiu analisar um grupo de satDNAs que possivelmente se originaram dentro da ordem Acipenseriformes, pois não foram encontrados em *Polypterus senegalus*, grupo externo da ordem. A origem pode ter ocorrido por formação *de novo* dessas sequências ou oriundas de outras sequências, fato que não foi possível descrever devido à baixa representatividade de membros da família Polypteridae, que contém 14 membros, mas apenas um foi estudado no presente estudo (Near; Thacker, 2024) e, da mesma forma, para a ordem Acipenseriformes, que possui uma boa representação de genomas em bases de dados públicas,

mas que carece de um completo sequenciamento de indivíduos deste grupo, o que poderia corroborar ou refutar a hipótese de origem desse grupo de satDNAs em *Acipenseriformes*, entretanto, os dados apresentados são consistentes com as hipóteses outrora apresentados por diversos autores, como Fry & Salser (1977) e Camacho et al., (2022) sobre origem e diversificação de satDNAs, e permite formular esta hipótese.

Uma hipótese que tem força é de que os satDNAs conservados no gênero *Acipenser* se originou há, pelo menos, 100 milhões de anos (Shen et al., 2020) pois não foram encontrados tanto no grupo externo do gênero, *Polyodon spathula*, como no grupo externo da Ordem, na família Polypteridae, e os padrões de expansão e contração das bibliotecas estão de acordo com a hipótese *Library* (Fry; Salser, 1977; Camacho et al., 2022; Šatović-Vukšić; Plohl, 2023), para confirmar esta hipótese é necessário expandir as análises, abarcando as demais espécies de *Acipenser* para traçar o caminho evolutivo destas sequências.

Na família conservada 5 observou-se que os satDNAs AoxSat01-4200 e AscSat01-4200 apresentam 98,571% de similaridade, o que levanta a hipótese para a conservação deste satDNA devido a uma possível função estrutural no genoma, como já foram reportados satDNAs abundantes que apresentam uma padrão de marcação centromérico/pericentromérico (vide PliSat01-225 e PliSat02-184 em Stornioli et al., 2021 e PmeSat05-247 e CmaSat05-247 em Goes et al., 2023) além do fato de que satDNAs são elementos importantes no centrômero (Hartley; O'Neill, 2019; Melters et al., 2013), porém, para validar, são necessário experimentos com Hibridação *in situ* de Fluorescência com sondas deste satDNAs para primeiro confirmar a sua localização citogenética e posteriormente estudos de expressão para analisar se, de fato, eles são expressos e podem apresentar uma função.

De forma a compreender se os satDNAs possuem alguma similaridade com elementos depositados na base de dados do NCBI uma análise de BLAST foi realizada, onde observou-se cinco resultados relevantes. Um satDNA conservado em algumas espécies de *Acipenser*, o *HindIII*, (Biltueva et al., 2020) foi encontrado no presente estudo, o AscSat02-171 (Figura 6) está presente em três das cinco espécies do estudo, sendo que em *Polyodon spathula* e *Acipenser oxyrinchus* ele não foi encontrado, enquanto que em *A. ruthenus* apresenta uma cobertura fragmentada, o que pode indicar que ele está sofrendo processo de divergência da sequência original e homogeneização (Šatović-Vukšić; Plohl, 2023), enquanto que em *A. schrenckii* e *A. sinensis* ele foi encontrado.

Em estudo previamente realizado, foi observado que o satDNA *HindIII* não foi encontrado em *P. spathula* e *A. oxyrinchus* e encontrado em *A. ruthenus* (Robles et al., 2004), o que o presente estudo corrobora, junto com o fato de que as sequências analisadas apresentaram baixa divergência, provavelmente pelo fato de que possuem uma evolução molecular mais lenta ou apresenta uma possível função (Ludwig et al., 2001) e, mesmo divergindo há aproximadamente 64 ma do ancestral em comum dos indivíduos analisados (Shedko, 2022) apresentam uma considerável taxa de conservação.

O satDNA AoxSat01-4200 foi encontrado em uma região de um possível *long non-coding RNA* (lncRNA), sendo um DNA satélite bem abundante em todas as espécies estudadas, o lncRNA possui 14,643 pares de base e o satDNA está presente apenas em uma região do lncRNA, o que levanta as hipóteses de que ou ele é ativamente transcrito pois possui uma função, visto que lncRNA são moléculas na qual apenas a região funcional é mais conservada, com as demais regiões menos reguladas (Zampetaki; Albrecht; Steinhofel, 2018), enquanto que a segunda hipótese é justamente que o satDNA está presente nessas regiões onde é passível que mutações ou inserções de sequências ocorram sem prejudicar o funcionamento do lncRNA num geral (Geisler; Coller, 2013; Zampetaki; Albrecht; Steinhofel, 2018).

Os demais casos são de moléculas presentes em possíveis regiões gênicas que podem ser expressas, em todos os casos os satDNAs estão presentes em regiões de íntrons ou na periferia dos genes, apesar da possível origem de um satDNAs ser muito debatida, é possível levantar a hipótese de que eles estão presentes nessas regiões devido a uma possível inserção via algum elemento transponível, já que existem evidências de que a movimentação de TEs pelo genoma pode carregar satDNAs, que, posteriormente, podem se expandir e gerar longos *arrays* (McGurk; Barbash, 2018; Šatović-Vukšić; Plohl, 2023). É importante destacar que, apesar de promissores, os resultados ainda precisam ser mais validados pois as sequências isoladas através da análise do NCBI são sequências previstas a partir do algoritmo e da característica dessas regiões dos *scaffolds*, sendo necessário primeiro a confirmação da existência dessa estrutura no genoma da(s) espécie(s) e, a posteriori, testar as hipóteses apresentadas.

6. Conclusão

Com o presente estudo foi possível descrever cinco novos satelitomas para a ordem Acipenseriformes e um para a família Polypteridae, contribuindo para a expansão da análise de satDNAs em Actinopterygii e sua evolução. Foi possível observar que estes elementos apresentam um padrão de conservação neste grupo, com uma quantia considerável dos elementos sendo conservados entre as espécies e indicando que certas sequências tenham se originado há, pelo menos, 100 milhões de anos, no gênero *Acipenser*, demonstrando que estas sequências podem ser conservadas por longos períodos, conforme já demonstrado em outros grupos e, agora, em Actinopterygii. Os resultados do BLAST permitiram concluir que o satDNA *HindIII* foi isolado, sendo o AscSat02-171, e corrobora o descrito previamente na literatura, enquanto outros satDNAs foram encontrados em possíveis regiões de mRNA e um deles em região de lncRNA, o que sugere uma possível transposição via elemento transponível.

7. Referências

BILTUEVA, Larisa S. *et al.* Genomic Organization and Physical Mapping of Tandemly Arranged Repetitive DNAs in Sterlet (*Acipenser ruthenus*). **Cytogenetic and Genome Research**, v. 152, n. 3, p. 148–157, 2017.

BILTUEVA, Larisa S. *et al.* Chromosome Distribution of Highly Conserved Tandemly Arranged Repetitive DNAs in the Siberian Sturgeon (*Acipenser baerii*). **Genes**, v. 11, n. 11, p. 1375, 20 nov. 2020.

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 1 ago. 2014.

BRONZI, P.; ROSENTHAL, H.; GESSNER, J. Global sturgeon aquaculture production: an overview. **Journal of Applied Ichthyology**, v. 27, n. 2, p. 169–175, 2011.

CAMACHO, Juan Pedro M. *et al.* Satellitome comparison of two oedipodine grasshoppers highlights the contingent nature of satellite DNA evolution. **BMC Biology**, v. 20, n. 1, p. 36, dez. 2022.

CHOUDHURY, Anindo; DICK, Terry A. The historical biogeography of sturgeons (Osteichthyes: Acipenseridae): a synthesis of phylogenetics, palaeontology and palaeogeography. **Journal of Biogeography**, v. 25, n. 4, p. 623–640, 1998.

DE LA HERRÁN, Roberto *et al.* Slow Rates of Evolution and Sequence Homogenization in an Ancient Satellite DNA Family of Sturgeons. **Molecular Biology and Evolution**, v. 18, n. 3, p. 432–436, 1 mar. 2001.

DE LIMA, Leonardo G.; RUIZ-RUANO, Francisco J. In-Depth Satellitome Analyses of 37 *Drosophila* Species Illuminate Repetitive DNA Evolution in the *Drosophila* Genus. **Genome Biology and Evolution**, v. 14, n. 5, p. evac064, 1 maio 2022.

DE MORAES, Renata Luiza Rosa *et al.* Chromosomal Rearrangements and Satellite DNAs: Extensive Chromosome Reshuffling and the Evolution of Neo-Sex Chromosomes in the Genus *Pyrrhulina* (Teleostei; Characiformes). **International Journal of Molecular Sciences**, v. 24, n. 17, p. 13654, jan. 2023.

DEHAL, Paramvir; BOORE, Jeffrey L. Two Rounds of Whole Genome Duplication in the Ancestral Vertebrate. **PLoS Biology**, v. 3, n. 10, p. e314, 6 set. 2005.

DOS SANTOS, Rodrigo Zeni *et al.* A Long-Term Conserved Satellite DNA That Remains Unexpanded in Several Genomes of Characiformes Fish Is Actively Transcribed. **Genome Biology and Evolution**, v. 13, n. 2, p. evab002, 1 fev. 2021.

DOS SANTOS, Rodrigo Zeni *et al.* Comparative satellite DNA mapping in species of the genus *Prochilodus* (Teleostei, Characiformes) and its evolutionary implications. **Genome**, v. 68, p. 1–8, jan. 2025.

DU, Kang *et al.* The sterlet sturgeon genome sequence and the mechanisms of segmental rediploidization. **Nature Ecology & Evolution**, v. 4, n. 6, p. 841–852, 30 mar. 2020.

DYSIN, Artem P. *et al.* Salmonidae Genome: Features, Evolutionary and Phylogenetic Characteristics. **Genes**, v. 13, n. 12, p. 2221, dez. 2022.

ELDER, John F.; TURNER, Bruce J. Concerted Evolution of Repetitive DNA Sequences in Eukaryotes. **The Quarterly Review of Biology**, v. 70, n. 3, p. 297–320, set. 1995.

FRIEDMAN, Matt. The early evolution of ray-finned fishes. **Palaeontology**, v. 58, n. 2, p. 213–228, 2015.

FRY, Kirk; SALSER, Winston. Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. **Cell**, v. 12, n. 4, p. 1069–1084, 1 dez. 1977.

GARRIDO-RAMOS, Manuel. Satellite DNA: An Evolving Topic. **Genes**, v. 8, n. 9, p. 230, 18 set. 2017.

GEISLER, Sarah; COLLIER, Jeff. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. **Nature Reviews Molecular Cell Biology**, v. 14, n. 11, p. 699–712, nov. 2013.

GOES, Caio Augusto Gomes *et al.* Revealing the Satellite DNA History in *Psalidodon* and *Astyanax* Characid Fish by Comparative Satellitomics. **Frontiers in Genetics**, v. 13, p. 884072, 21 jun. 2022.

GOES, Caio Augusto Gomes *et al.* The Satellite DNA Catalogues of Two Serrasalminae (Teleostei, Characiformes): Conservation of General satDNA Features over 30 Million Years. **Genes**, v. 14, n. 1, p. 91, jan. 2023.

HARTLEY, Gabrielle; O'NEILL, Rachel J. Centromere Repeats: Hidden Gems of the Genome. **Genes**, v. 10, n. 3, p. 223, mar. 2019.

HUGHES, Lily C. *et al.* Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. **Proceedings of the National Academy of Sciences**, v. 115, n. 24, p. 6249–6254, 12 jun. 2018.

KOJIMA, Kenji K. Structural and sequence diversity of eukaryotic transposable elements. **Genes & Genetic Systems**, v. 94, n. 6, p. 233–252, 2019.

KRETSCHMER, Rafael *et al.* Satellitome analysis illuminates the evolution of ZW sex chromosomes of Triportheidae fishes (Teleostei: Characiformes). **Chromosoma**, v. 131, n. 1–2, p. 29–45, jun. 2022.

LANFREDI, M. *et al.* Chromosomal location and evolution of a satellite DNA family in seven sturgeon species. **Chromosome Research**, v. 9, n. 1, p. 47–52, 1 jan. 2001.

LÉVÊQUE, Christian. Biodiversity dynamics and conservation : the freshwater fish of tropical Africa. [S.d.].

LUDWIG, A. *et al.* Genome duplication events and functional reduction of ploidy levels in sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). **Genetics**, v. 158, n. 3, p. 1203–1215, jul. 2001.

LUDWIG, A. Identification of Acipenseriformes species in trade. **Journal of Applied Ichthyology**, v. 24, n. s1, p. 2–19, 2008.

MCGURK, Michael P.; BARBASH, Daniel A. Double insertion of transposable elements provides a substrate for the evolution of satellite DNA. **Genome Research**, v. 28, n. 5, p. 714–725, 5 jan. 2018.

MELTERS, Daniël P. *et al.* Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. **Genome Biology**, v. 14, n. 1, p. R10, 30 jan. 2013.

MORA, Pablo *et al.* Satellitome Analysis of *Adalia bipunctata* (Coleoptera): Revealing Centromeric Turnover and Potential Chromosome Rearrangements in a Comparative Interspecific Study. **International Journal of Molecular Sciences**, v. 25, n. 17, p. 9214, jan. 2024.

NEAR, Thomas J.; THACKER, Christine E. Phylogenetic Classification of Living and Fossil Ray-Finned Fishes (Actinopterygii). **Bulletin of the Peabody Museum of Natural History**, v. 65, n. 1, p. 3–302, abr. 2024.

NOVÁK, Petr; NEUMANN, Pavel; MACAS, Jiří. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. **Nature Protocols**, v. 15, n. 11, p. 3745–3776, nov. 2020.

PIKITCH, Ellen K. *et al.* Status, trends and management of sturgeon and paddlefish fisheries. **Fish and Fisheries**, v. 6, n. 3, p. 233–265, 2005.

RAPOSO, António *et al.* Eating Sturgeon: An Endangered Delicacy. **Sustainability**, v. 15, n. 4, p. 3511, jan. 2023.

RIESEBERG, Loren H. *et al.* Hybridization and the colonization of novel habitats by annual sunflowers. **Genetica**, v. 129, n. 2, p. 149–165, 1 fev. 2007.

ROBLES, Francisca *et al.* Evolution of ancient satellite DNAs in sturgeon genomes. **Gene**, v. 338, n. 1, p. 133–142, ago. 2004.

RUIZ-RUANO, Francisco J. *et al.* High-throughput analysis of the satellitome illuminates satellite DNA evolution. **Scientific Reports**, v. 6, n. 1, p. 28333, 7 jul. 2016.

SALES-OLIVEIRA, Vanessa C. *et al.* Evolution of ancient satellite DNAs in extant alligators and caimans (Crocodylia, Reptilia). **BMC Biology**, v. 22, n. 1, p. 47, 27 fev. 2024.

ŠATOVIĆ-VUKŠIĆ, Eva; PLOHL, Miroslav. Satellite DNAs—From Localized to Highly Dispersed Genome Components. **Genes**, v. 14, n. 3, p. 742, 17 mar. 2023.

SAYERS, Eric W. *et al.* Database resources of the National Center for Biotechnology Information in 2025. **Nucleic Acids Research**, v. 53, n. D1, p. D20–D29, 6 jan. 2025.

SCACCHETTI, Priscilla Cardim *et al.* Repetitive DNA Sequences and Evolution of ZZ/ZW Sex Chromosomes in *Characidium* (Teleostei: Characiformes). **PLOS ONE**, v. 10, n. 9, p. e0137231, 15 set. 2015.

SCHMIEDER, Robert; EDWARDS, Robert. Fast Identification and Removal of Sequence Contamination from Genomic and Metagenomic Datasets. **PLOS ONE**, v. 6, n. 3, p. e17288, 9 mar. 2011.

SÉMON, Marie; WOLFE, Kenneth H. Reciprocal gene loss between *Tetraodon* and zebrafish after whole genome duplication in their ancestor. **Trends in Genetics**, v. 23, n. 3, p. 108–112, 1 mar. 2007.

SERRANO-FREITAS, Érica A. *et al.* Satellite DNA content of B chromosomes in the characid fish *Characidium gomesi* supports their origin from sex chromosomes. **Molecular Genetics and Genomics**, v. 295, n. 1, p. 195–207, 1 jan. 2020.

SETTI, Príncia Grejo *et al.* Evolution of bird sex chromosomes: a cytogenomic approach in Palaeognathae species. **BMC Ecology and Evolution**, v. 24, n. 1, p. 51, 23 abr. 2024.

SHEDKO, S. V. Molecular Dating of Phylogeny of Sturgeons (Acipenseridae) Based on Total Evidence Analysis. **Russian Journal of Genetics**, v. 58, n. 6, p. 718–729, jun. 2022.

SHEN, Yanjun *et al.* Phylogenetic perspective on the relationships and evolutionary history of the Acipenseriformes. **Genomics**, v. 112, n. 5, p. 3511–3517, set. 2020.

SILVA, Duílio M. Z. de A. *et al.* High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus *Astyanax*. **Scientific Reports**, v. 7, n. 1, p. 12726, 5 out. 2017.

SMIT, A. F. A.; HUBLEY, R.; GREEN, P. **RepeatMasker Open-4.0**. Disponível em: <<https://www.repeatmasker.org/>>. Acesso em: 19 fev. 2025.

SMITH, George P. Evolution of Repeated DNA Sequences by Unequal Crossover. **Science**, v. 191, n. 4227, p. 528–535, 13 fev. 1976.

STEPHAN, W.; CHO, S. Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. **Genetics**, v. 136, n. 1, p. 333–341, 1 jan. 1994.

STORNIOLI, José Henrique Forte *et al.* The B Chromosomes of *Prochilodus lineatus* (Teleostei, Characiformes) Are Highly Enriched in Satellite DNAs. **Cells**, v. 10, n. 6, p. 1527, jun. 2021.

THE GALAXY COMMUNITY. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. **Nucleic Acids Research**, v. 52, n. W1, p. W83–W94, 5 jul. 2024.

TRIFONOV, Vladimir A. *et al.* Evolutionary plasticity of Acipenseriform genomes. **Chromosoma**, v. 125, n. 4, p. 661–668, 1 set. 2016.

UTSUNOMIA, Ricardo *et al.* A Glimpse into the Satellite DNA Library in Characidae Fish (Teleostei, Characiformes). **Frontiers in Genetics**, v. 8, 14 ago. 2017.

VAN DE PEER, Yves; MAERE, Steven; MEYER, Axel. The evolutionary significance of ancient genome duplications. **Nature Reviews Genetics**, v. 10, n. 10, p. 725–732, out. 2009.

VAN DER LAAN, Richard; ESCHMEYER, William N.; FRICKE, Ronald. Family-group names of Recent fishes. **Zootaxa**, v. 3882, n. 1, 11 nov. 2014.

VANDEPOELE, Klaas *et al.* Major events in the genome evolution of vertebrates: Paraneome age and size differ considerably between ray-finned fishes and land vertebrates. **Proceedings of the National Academy of Sciences**, v. 101, n. 6, p. 1638–1643, 10 fev. 2004.

YUAN, Zihao *et al.* Comparative genome analysis of 52 fish species suggests differential associations of repetitive elements with their living aquatic environments. **BMC Genomics**, v. 19, n. 1, p. 141, dez. 2018.

ZAMPETAKI, Anna; ALBRECHT, Andreas; STEINHOFEL, Kathleen. Long Non-coding RNA Structure and Function: Is There a Link? **Frontiers in Physiology**, v. 9, 24 ago. 2018.

8. Material Suplementar

Tabela S1. Tabela completa do satelitoma de *Polyodon spathula*.

SF	Família de satDNA	Monômero	Abundância	Divergência	A+T (%)
	PspSat01-219	219	0.002562666	1.51	58
	PspSat02-2520	2520	0.001994443	8.26	62
	PspSat03-2182	2182	0.001299179	3.61	63
2	PspSat04-178	178	0.001200525	9.99	62
	PspSat05-80	80	0.001199661	9.66	43
	PspSat06-69	69	0.001101885	3.69	59
2	PspSat07-178	178	0.000994863	7	60
	PspSat08-1679	1679	0.000805663	9.93	55
	PspSat09-29	29	0.000741369	6.49	55
	PspSat10-18	18	0.000729849	11.84	55
	PspSat11-359	359	0.000693656	5.77	60
3	PspSat12-138	138	0.000664864	15.71	46
	PspSat13-47	47	0.000634725	3.91	55
1	PspSat14-497	497	0.000523954	2.15	46
1	PspSat15-263	263	0.00051763	7.71	45
2	PspSat16-177	177	0.000515587	6.73	61
2	PspSat17-177	177	0.000500674	11.83	67
1	PspSat18-338	338	0.000484559	3.35	47
	PspSat19-39	39	0.000441413	6.76	48
	PspSat20-969	969	0.000438125	7.62	37
	PspSat21-56	56	0.00040048	9.33	57
2	PspSat22-357	357	0.000379581	3.42	62
	PspSat23-57	57	0.000369864	3.86	61
1	PspSat24-587	587	0.000356037	2.67	46
3	PspSat25-315	315	0.00034075	5.79	45
	PspSat26-36	36	0.000340086	4.21	55
3	PspSat27-496	496	0.000320589	6.94	45
1	PspSat28-290	290	0.000288768	6.43	44
	PspSat29-510	510	0.000278003	8.18	59
2	PspSat30-177	177	0.000274088	13.8	62
	PspSat31-224	224	0.000253507	11.04	57
4	PspSat32-814	814	0.000252144	1.55	56
3	PspSat33-1334	1334	0.000248478	10.57	43
4	PspSat34-490	490	0.000247535	5.52	55
4	PspSat35-1293	1293	0.00023869	2.69	54
4	PspSat36-550	550	0.000229485	3.54	56
1	PspSat37-150	150	0.000219338	9.38	42
2	PspSat38-178	178	0.00019542	7.9	61
1	PspSat39-1346	1346	0.000193799	1.64	45

	PspSat40-46	46	0.000191329	2.62	65
2	PspSat41-358	358	0.000187695	3.91	59
	PspSat42-58	58	0.000185541	6.87	55
	PspSat43-324	324	0.000178033	8.83	60
	PspSat44-82	82	0.000171842	1.89	57
	PspSat45-39	39	0.000171741	5.47	61
	PspSat46-43	43	0.000169019	6.68	65
4	PspSat47-518	518	0.000158206	4.98	58
	PspSat48-30	30	0.00015448	5.94	73
	PspSat49-47	47	0.000152497	9.29	48
	PspSat50-40	40	0.00015067	3.45	60
	PspSat51-65	65	0.000144235	5.22	61
	PspSat52-27	27	0.000143377	10.9	62
	PspSat53-631	631	0.000137894	7.26	52
1	PspSat54-188	188	0.000136512	8.06	42
	PspSat55-29	29	0.000131607	6.21	62
	PspSat56-68	68	0.000131535	8.79	57
	PspSat57-1006	1006	0.000131215	1.81	54
	PspSat58-49	49	0.000126357	7.66	42
	PspSat59-40	40	0.000122782	5.23	65
4	PspSat60-441	441	0.000114841	8.74	54
	PspSat61-32	32	0.000112509	5.19	53
	PspSat62-24	24	0.000110237	14.69	37
	PspSat63-40	40	8.95E-05	2.12	50
4	PspSat64-287	287	8.61E-05	7.17	55
	PspSat65-47	47	8.36E-05	3.98	74

Tabela S2. Tabela completa do satelitoma de *Acipenser oxyrinchus*.

SF	Família de satDNA	Monômero	Abundância	Divergência	A+T (%)
	AoxSat01-4200	4200	0.010950727	2.92	53
	AoxSat02-2596	2596	0.010918123	4.37	62
	AoxSat03-25	25	0.005136328	20.91	56
1	AoxSat04-1303	1303	0.003768878	3.27	49
	AoxSat05-40	40	0.002833053	5.75	55
1	AoxSat06-473	473	0.002084321	2.1	40
2	AoxSat07-297	297	0.000765766	13.84	62
	AoxSat08-358	358	0.000732273	1.23	58
	AoxSat09-241	241	0.000549446	3.66	46
	AoxSat10-248	248	0.000513331	4.39	52
	AoxSat11-187	187	0.000461855	5.5	59
	AoxSat12-42	42	0.000412635	8.86	50
	AoxSat13-26	26	0.000375486	12.27	38
	AoxSat14-52	52	0.000254623	6.98	60
	AoxSat15-138	138	0.000226222	9.61	57
	AoxSat16-80	80	0.000208188	6.32	58
	AoxSat17-52	52	0.000206629	9.46	44
	AoxSat18-60	60	0.00019253	6.63	55
	AoxSat19-26	26	0.000190101	14.42	54
	AoxSat20-31	31	0.000180475	4.94	52
	AoxSat21-57	57	0.000176132	3.99	60
2	AoxSat22-93	93	0.00016436	9.76	55
	AoxSat23-178	178	0.00016149	15.2	59
	AoxSat24-30	30	0.000157274	10.7	43
	AoxSat25-34	34	0.000149581	7.51	50
	AoxSat26-31	31	0.000142392	7.81	48
	AoxSat27-42	42	0.000135483	11.52	36
	AoxSat28-33	33	0.000119114	5.15	55
	AoxSat29-51	51	0.000113078	3.06	55
	AoxSat30-30	30	0.00010563	7.2	57
	AoxSat31-35	35	0.000102754	8.23	51
	AoxSat32-44	44	9.58354E-05	8.45	48
	AoxSat33-22	22	8.98514E-05	7.7	59
	AoxSat34-17	17	8.15728E-05	12.92	47

Tabela S3. Tabela completa do satelitoma de *Acipenser schrenckii*.

SF	Família de satDNA	Monômero	Abundância	Divergência	A+T (%)
	AscSat01-4200	4200	0.009556292	2.82	53
	AscSat02-171	171	0.009548315	4.43	65
	AscSat03-656	656	0.003879089	7.41	58
	AscSat04-952	952	0.003659897	7.92	62
	AscSat05-2441	2441	0.001788477	6.88	65
3	AscSat06-40	40	0.001617712	8.23	55
	AscSat07-271	271	0.001417546	3.86	55
	AscSat08-25	25	0.001252631	11.73	44
	AscSat09-513	513	0.001161345	1.89	52
4	AscSat10-135	135	0.001063056	10.91	48
4	AscSat11-36	36	0.000982894	9.51	38
	AscSat12-89	89	0.000733309	11.63	61
4	AscSat13-188	188	0.000571939	4.69	48
1	AscSat14-179	179	0.000524619	9.12	59
	AscSat15-982	982	0.000517029	4.96	66
	AscSat16-1291	1291	0.000472299	7.16	58
3	AscSat17-40	40	0.000455953	15.66	57
1	AscSat18-178	178	0.000312517	9.63	60
	AscSat19-625	625	0.000303255	8.54	40
	AscSat20-52	52	0.00029812	7.44	60
2	AscSat21-27	27	0.000274576	12.01	48
	AscSat22-242	242	0.000262499	9.52	43
	AscSat23-1600	1600	0.000247828	1.04	66
	AscSat24-1488	1488	0.000229729	6.6	56
	AscSat25-57	57	0.000196814	4.8	60
	AscSat26-28	28	0.000196129	5.01	53
	AscSat27-43	43	0.000195324	7.37	65
	AscSat28-80	80	0.000187686	6.41	57
	AscSat29-121	121	0.000183261	7	53
	AscSat30-59	59	0.000158045	6.91	56
	AscSat31-65	65	0.000157986	6.88	55
3	AscSat32-481	481	0.000139455	17.45	60
2	AscSat33-33	33	0.000131054	6.25	45
	AscSat34-22	22	0.000128852	12.12	50
	AscSat35-49	49	0.000117904	8.67	55
	AscSat36-645	645	0.000100474	4.02	59
	AscSat37-40	40	7.72E-05	5.41	60
	AscSat38-31	31	9.77E-06	5.36	51
	AscSat39-179	179	9.67E-06	8.86	65
	AscSat40-51	51	8.95E-06	1.83	72
	AscSat41-187	187	8.82E-06	3.04	59

	AscSat42-185	185	7.30E-06	5.4	61
--	--------------	-----	----------	-----	----

Tabela S4. Tabela completa do satelitoma de *Acipenser sinensis*.

SF	Família de satDNA	Monômero	Abundância	Divergência	A+T (%)
	AsiSat01-271	271	0.003701484	3.82	54
	AsiSat02-789	789	0.003270578	2.33	57
4	AsiSat03-25	25	0.002650693	11.19	44
	AsiSat04-188	188	0.00135224	5.45	47
2	AsiSat05-28	28	0.001025199	6.01	53
	AsiSat06-22	22	0.000821329	4.79	45
	AsiSat07-26	26	0.000773713	3.97	50
	AsiSat08-47	47	0.000675442	6.41	48
	AsiSat09-55	55	0.000611604	3.52	52
3	AsiSat10-33	33	0.000586215	10.29	39
	AsiSat11-36	36	0.000548691	9.64	38
5	AsiSat12-43	43	0.000501359	3.5	46
	AsiSat13-145	145	0.000485588	6.26	36
1	AsiSat14-29	29	0.000410863	5.53	48
	AsiSat15-47	47	0.000405411	8.38	57
	AsiSat16-36	36	0.000390281	6.29	44
	AsiSat17-241	241	0.000386593	9.69	43
	AsiSat18-46	46	0.000339841	10.57	45
	AsiSat19-40	40	0.000294349	6.38	45
	AsiSat20-34	34	0.000292901	11.43	41
	AsiSat21-171	171	0.000268798	5.85	65
	AsiSat22-1488	1488	0.000267466	4.12	63
	AsiSat23-39	39	0.000263824	3.97	38
1	AsiSat24-21	21	0.00023825	13.67	47
	AsiSat25-28	28	0.000225225	7.55	53
	AsiSat26-39	39	0.000222604	12.04	35
	AsiSat27-49	49	0.00021306	8.08	44
	AsiSat28-179	179	0.000171306	9.41	56
	AsiSat29-40	40	0.000168146	5.81	52
4	AsiSat30-25	25	0.000163519	15.59	56
	AsiSat31-43	43	0.000159359	3.74	46
	AsiSat32-36	36	0.000159027	6.76	47
	AsiSat33-56	56	0.000155546	7.55	44
	AsiSat34-41	41	0.000153007	9.28	41
	AsiSat35-45	45	0.000143714	5.41	53
	AsiSat36-30	30	0.000133382	9.36	46
	AsiSat37-18	18	0.00013096	15.63	55
2	AsiSat38-22	22	0.00013072	13.29	59
	AsiSat39-38	38	0.000130352	7.93	50
2	AsiSat40-28	28	0.000120345	6.88	50

	AsiSat41-34	34	0.000115665	6.17	47
	AsiSat42-30	30	0.000108162	3.1	50
3	AsiSat43-37	37	0.000107379	5.64	40
	AsiSat44-69	69	0.000107274	3.3	46
	AsiSat45-48	48	9.77E-05	3.64	45
	AsiSat46-37	37	9.72E-06	4.41	45
	AsiSat47-39	39	9.52E-06	3.37	43
	AsiSat48-50	50	9.46E-06	2.02	52
	AsiSat49-27	27	8.97E-06	10.34	48
	AsiSat50-49	49	8.67E-06	6.96	51
	AsiSat51-29	29	8.31E-06	8.35	48
	AsiSat52-38	38	7.38E-06	6.65	42
5	AsiSat53-47	47	6.21E-06	4.87	44

Tabela S5. Satelitoma completo de *Acipenser ruthenus*.

SF	Família de satDNA	Monômero	Abundância	Divergência	A+T (%)
	AruSat01-802	802	0.004207056	9.57	58
	AruSat02-26	26	0.003838254	10.84	42
	AruSat03-579	579	0.002077354	17.69	60
	AruSat04-3303	3303	0.001348726	6.73	59
	AruSat05-2441	2441	0.000526176	3.44	59
	AruSat06-43	43	0.000489687	6.56	53
	AruSat07-46	46	0.000483128	7.79	50
	AruSat08-31	31	0.000443658	4.99	51
	AruSat09-167	167	0.000434438	2.03	66
	AruSat10-152	152	0.000432766	6.26	36
	AruSat11-138	138	0.000388734	5.84	55
	AruSat12-36	36	0.000339815	5.14	50
	AruSat13-39	39	0.000330842	5.37	48
	AruSat14-121	121	0.000310889	6.4	48
	AruSat15-224	224	0.000277489	3.32	44
	AruSat16-82	82	0.000259385	5.08	48
	AruSat17-49	49	0.000253127	4.16	48
	AruSat18-1149	1149	0.000227071	4.62	49
	AruSat19-40	40	0.00022257	4.46	60
	AruSat20-1059	1059	0.000220009	2.57	65
	AruSat21-68	68	0.000201911	3.93	39
	AruSat22-38	38	0.000201887	8.07	42
	AruSat23-178	178	0.000188131	6.95	55
1	AruSat24-50	50	0.00018539	3.29	56
	AruSat25-24	24	0.000162015	9.08	41
	AruSat26-26	26	0.000143368	5.39	46
	AruSat27-32	32	0.00010367	5.56	53
1	AruSat28-179	179	0.000101324	11.51	63

	AruSat29-41	41	9.83E-05	4.96	46
	AruSat30-36	36	8.04E-05	3.28	58

Tabela S6. Satelitoma completo de *Polypterus senegalus*.

SF	Família de satDNA	Monômero	Abundância	Divergência	A+T (%)
	PseSat01-485	485	0.005085299	5.19	54
	PseSat02-38	38	0.000630884	13.47	57
	PseSat03-1345	1345	0.000269918	5.74	61
	PseSat04-1017	1017	0.000221878	1.57	58
	PseSat05-94	94	0.000179339	6.27	53
	PseSat06-940	940	0.000160463	7.92	57
	PseSat07-388	388	0.00013277	4.01	60
	PseSat08-1206	1206	0.000119162	1.07	54
	PseSat09-142	142	6.43E-05	5.61	65

Tabela S7. Satélites retirados da análise de presença em todas as espécies por possuírem mais de 80% de similaridade. * Satélites mantidos como representantes do respectivo agrupamento e mantidos na análise.

SatDNA I *	SatDNA II	SatDNA III	SatDNA IV
AcaSat12-21	AsiSat38-22		
AoxSat03-25	AsiSat30-25		
AruSat02-26	AscSat08-25	AsiSat03-25	
AoxSat05-40	AscSat06-40	AscSat17-40	
AscSat11-36	AsiSat11-36		
AoxSat09-241	AscSat22-242	AsiSat17-241	
AscSat41-187	AoxSat11-187		
AruSat07-46	AoxSat12-42		
AoxSat14-52	AscSat20-52		
AoxSat15-138	AruSat11-138		
AscSat07-271	AsiSat01-271		
AscSat28-80	AoxSat16-80		
AscSat31-65	AoxSat18-60		
AruSat08-31	AscSat38-31	AoxSat20-31	
AoxSat21-57	AscSat25-57	PspSat23-57	
AoxSat23-178	AruSat23-178	AscSat14-179	AsiSat28-179
AruSat09-167	AscSat02-171	AsiSat21-171	
AruSat30-36	PspSat26-36		
AscSat33-33	AsiSat10-33		
AscSat26-28	AsiSat25-28		
AscSat27-43	PspSat46-43		
AscSat30-59	PspSat42-58		
AscSat35-49	PspSat49-47		
AsiSat29-40	PspSat63-40		
AoxSat01-4200	AscSat01-4200		

Tabela S8. Análise comparativa dos membros da ordem Acipenseriformes, todos os satélites apresentam no mínimo 50% de similaridade.

Família Conservada	satDNA:	Família Conservada	satDNA:	Família Conservada	satDNA:
1	AruSat20-1059	7	AoxSat05-40	18	AoxSat23-178
	AscSat23-1600		AruSat03-579		AruSat23-178
2	AoxSat31-35		AscSat06-40		AruSat28-179
	AruSat27-32		AscSat17-40		AscSat14-179
	AruSat29-41		AscSat32-481		AscSat18-178
	AscSat34-22	8	AoxSat22-93		AsiSat28-179
	AsiSat05-28		AscSat10-135		PspSat04-178
	AsiSat14-29		AscSat11-36		PspSat07-178
	AsiSat16-36		AscSat13-188		PspSat16-177
	AsiSat24-21		AsiSat11-36		PspSat17-177
	AsiSat36-30	9	AoxSat09-241		PspSat22-357
	AsiSat38-22		AruSat15-224		PspSat30-177
	AsiSat40-28		AscSat22-242		PspSat38-178
	PspSat55-29		AsiSat17-241		PspSat41-358
	PspSat62-24	10	AscSat41-187	19	AoxSat24-30
3	AoxSat04-1303		AoxSat11-187		AsiSat24-21
	AoxSat06-473	11	AruSat07-46	20	AruSat09-167
	AoxSat08-358		AoxSat12-42		AscSat02-171
	AruSat10-152	12	AoxSat14-52		AsiSat21-171
	AscSat19-625		AscSat20-52	21	AruSat30-36
	AsiSat13-145	13	AoxSat15-138		AsiSat41-34
	PspSat08-1679		AruSat11-138		PspSat26-36

	PspSat11-359		AscSat07-271	22	AsiSat22-1488
4	AoxSat13-26		AsiSat01-271		AscSat15-982
	AoxSat33-22	14	AscSat28-80	23	AscSat21-27
	AsiSat38-22		AoxSat16-80		AscSat33-33
	AsiSat40-28	15	AscSat31-65		AsiSat10-33
5	AoxSat01-4200		AoxSat18-60		AsiSat43-37
	AscSat01-4200	16	AruSat08-31	24	AscSat26-28
6	AoxSat03-25		AscSat38-31		AsiSat25-28
	AruSat02-26		AoxSat20-31	25	AscSat27-43
	AscSat08-25	17	AoxSat21-57		PspSat46-43
	AsiSat03-25		AscSat25-57	26	AscSat30-59
	AsiSat30-25		PspSat23-57		PspSat42-58
				27	AscSat35-49
					PspSat49-47
				28	AsiSat29-40
					PspSat63-40

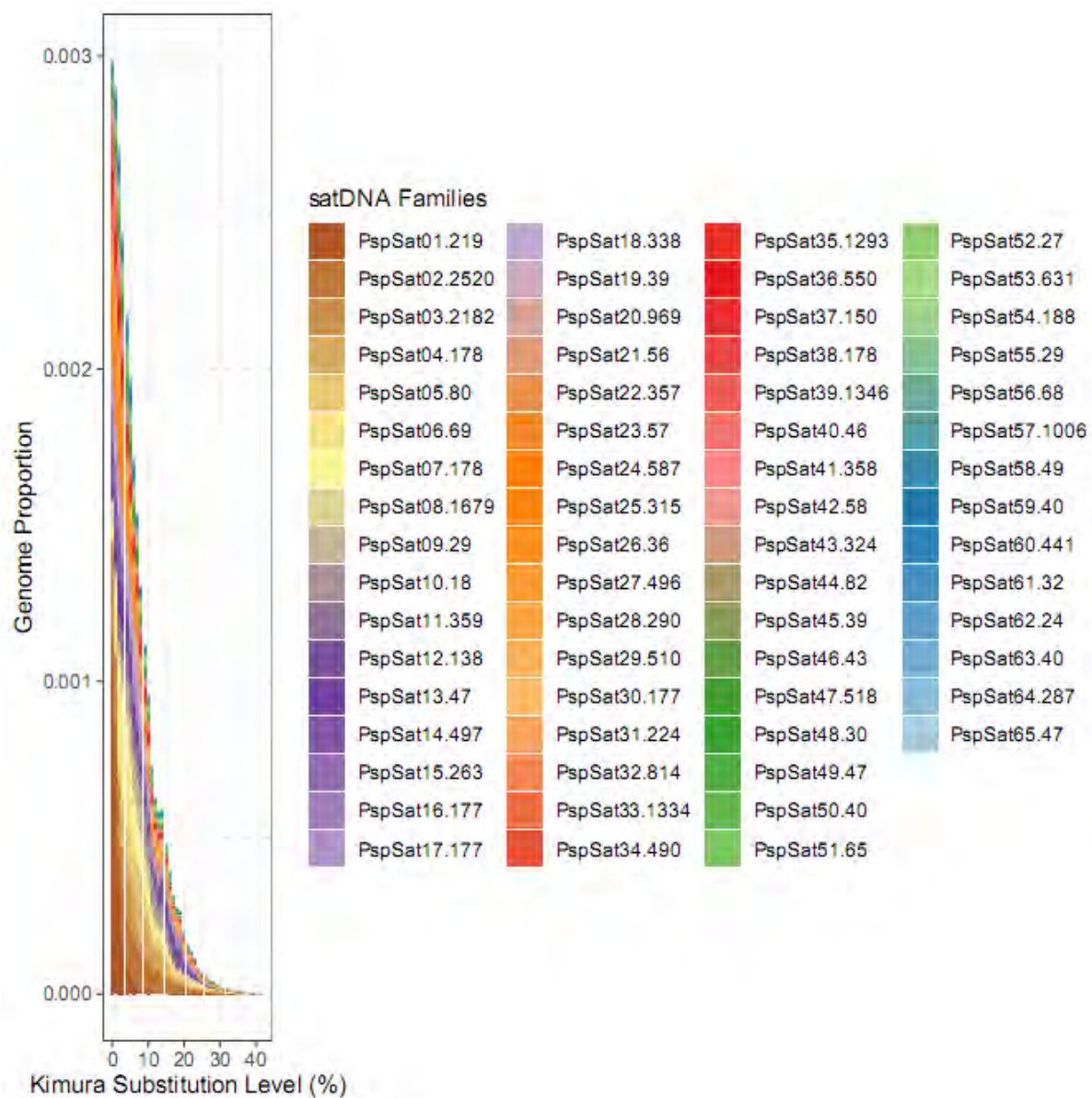


Figura S1. RepeatLandscape de *Polyodon spathula*.

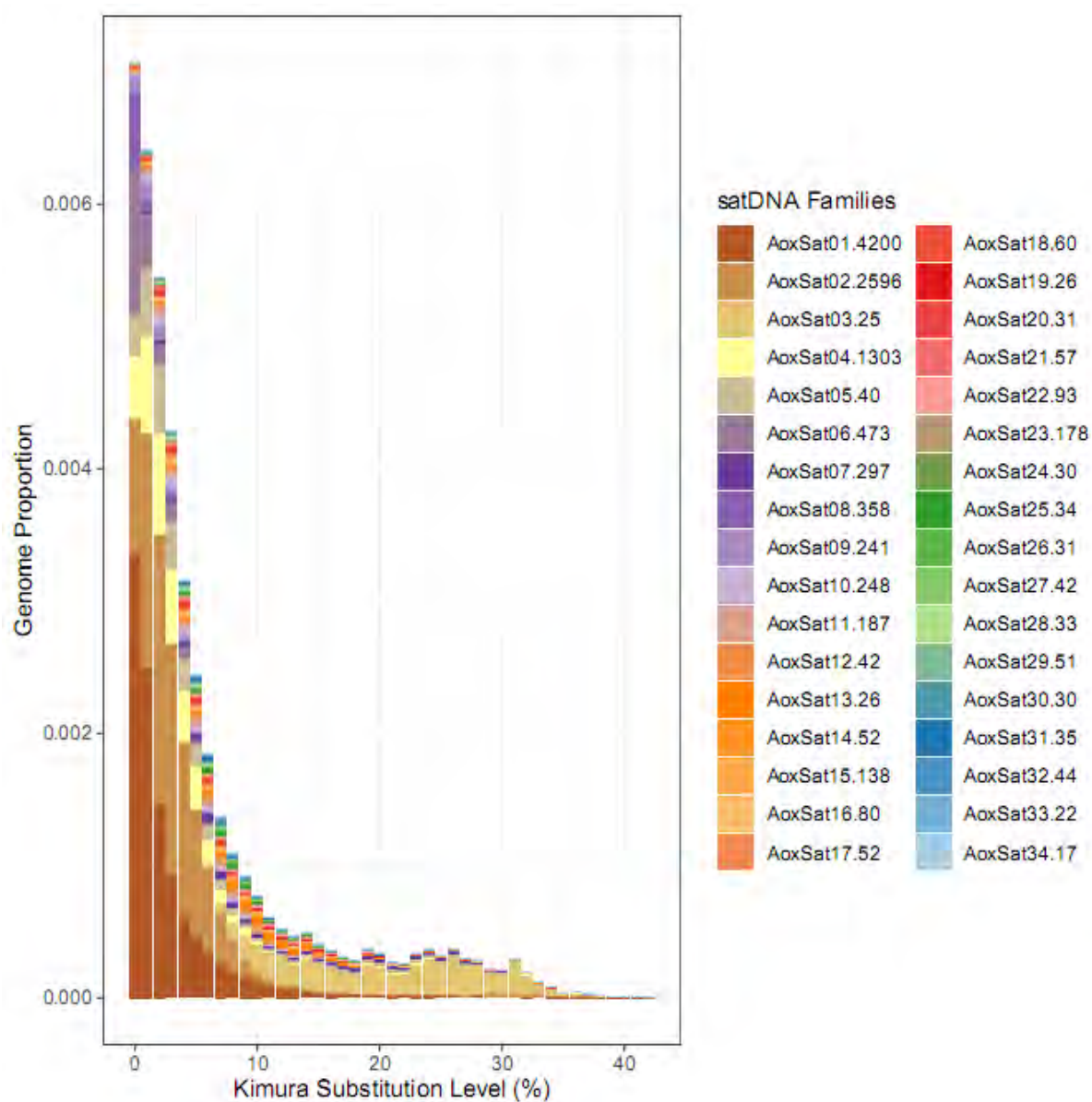


Figura S2. RepeatLandscape de *Acipenser oxyrinchus*.

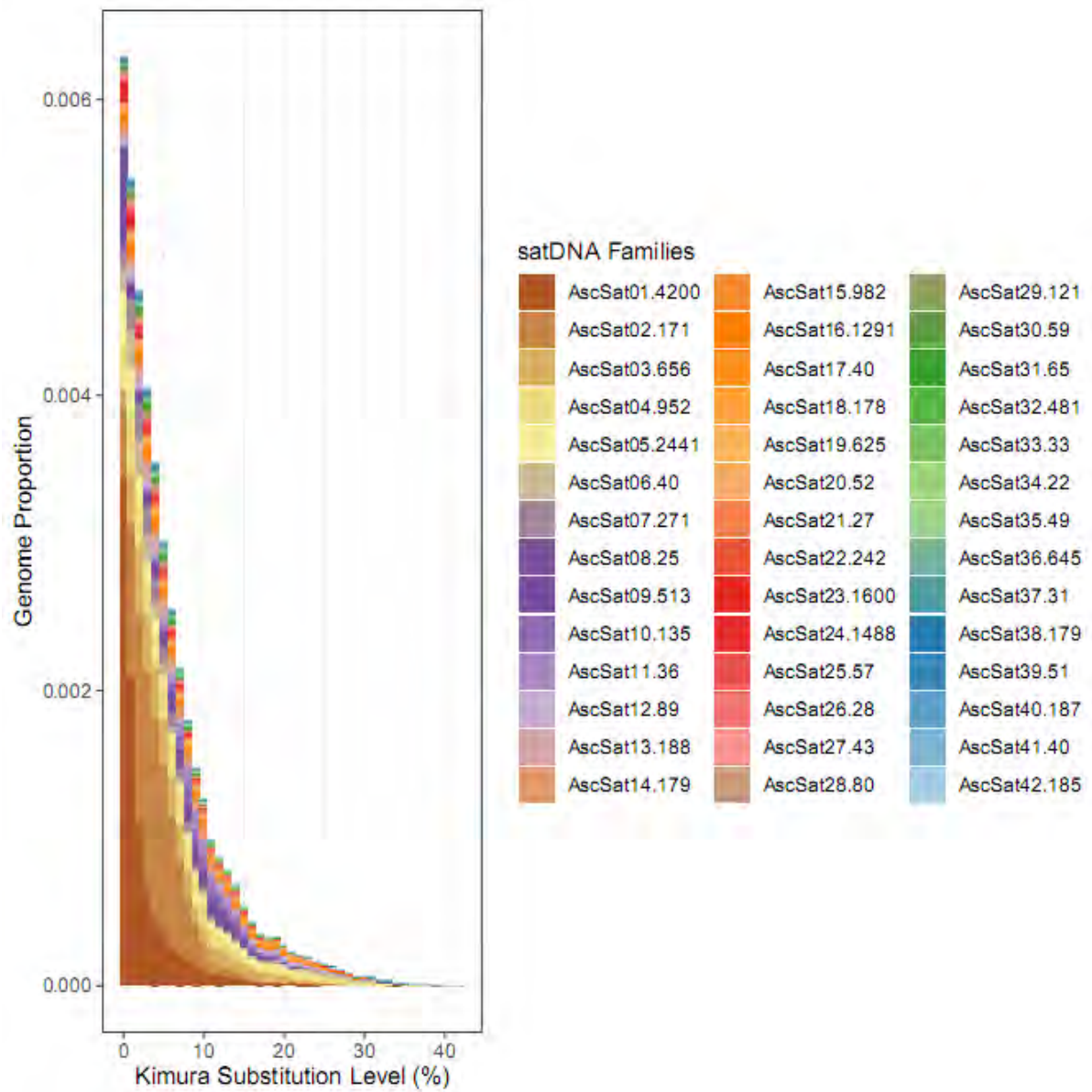


Figura S3. RepeatLandscape de *Acipenser schrenckii*.

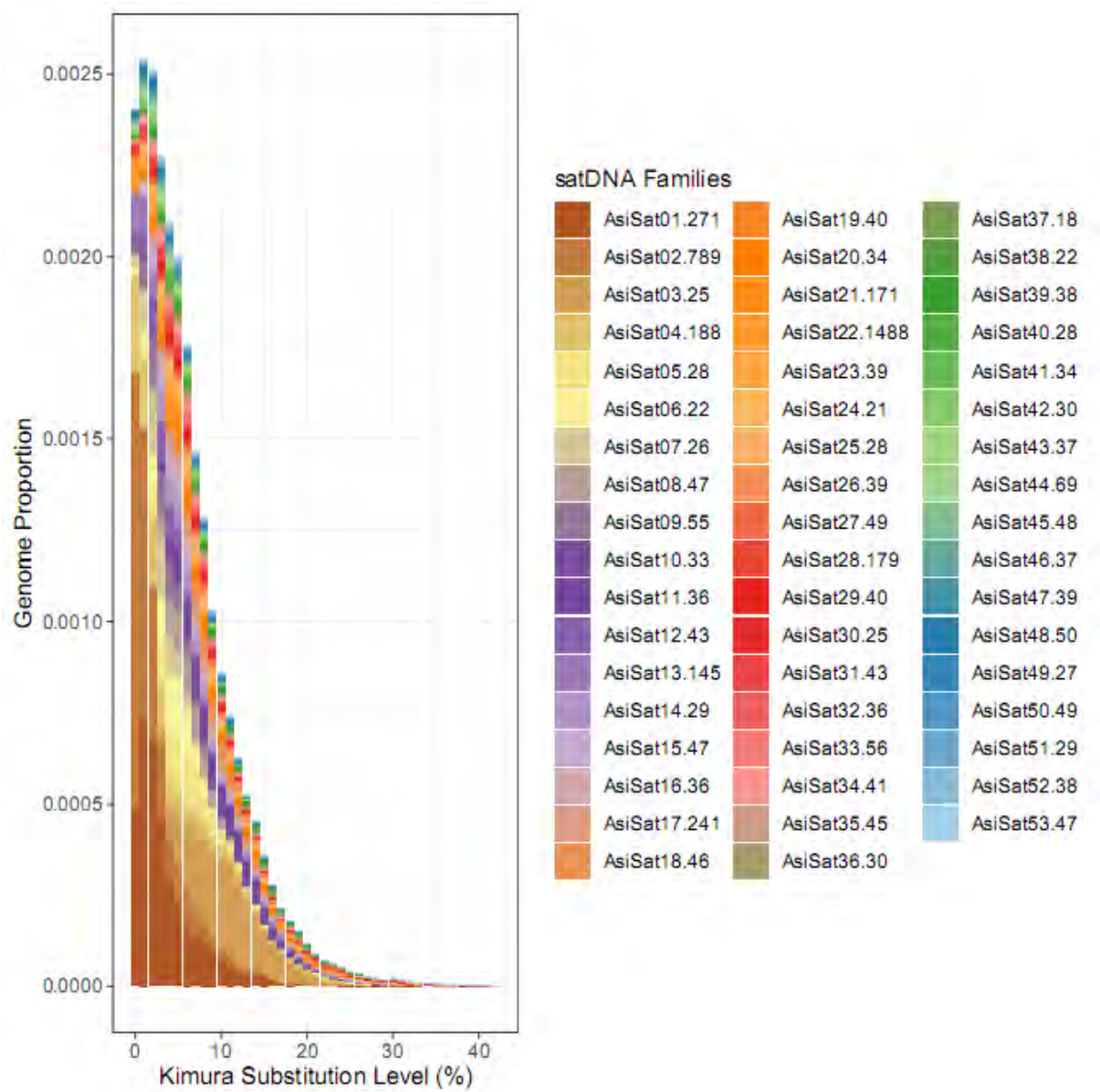


Figura S4. RepeatLandscape de *Acipenser sinensis*.

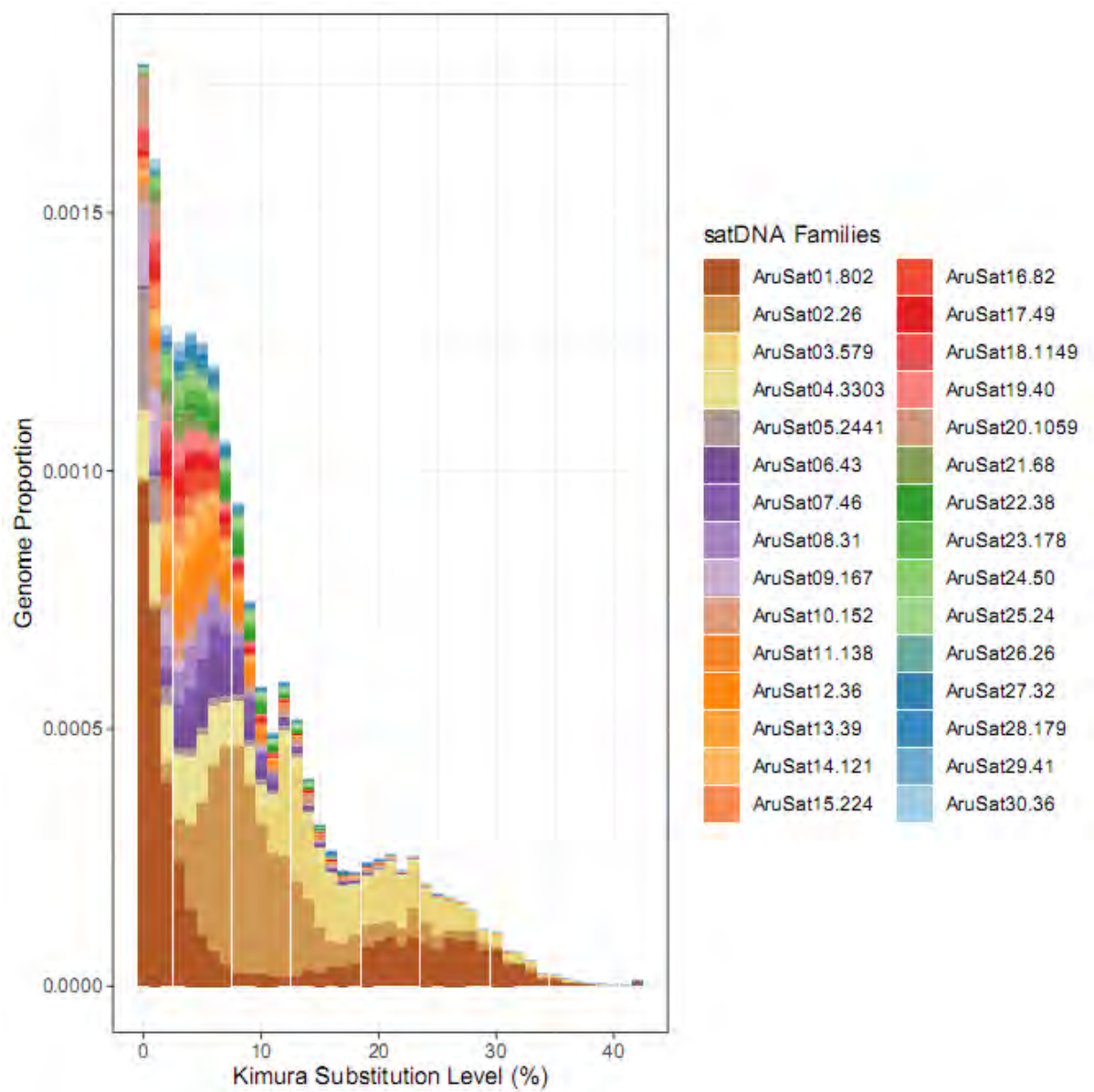


Figura S5. RepeatLandscape de *Acipenser ruthenus*.

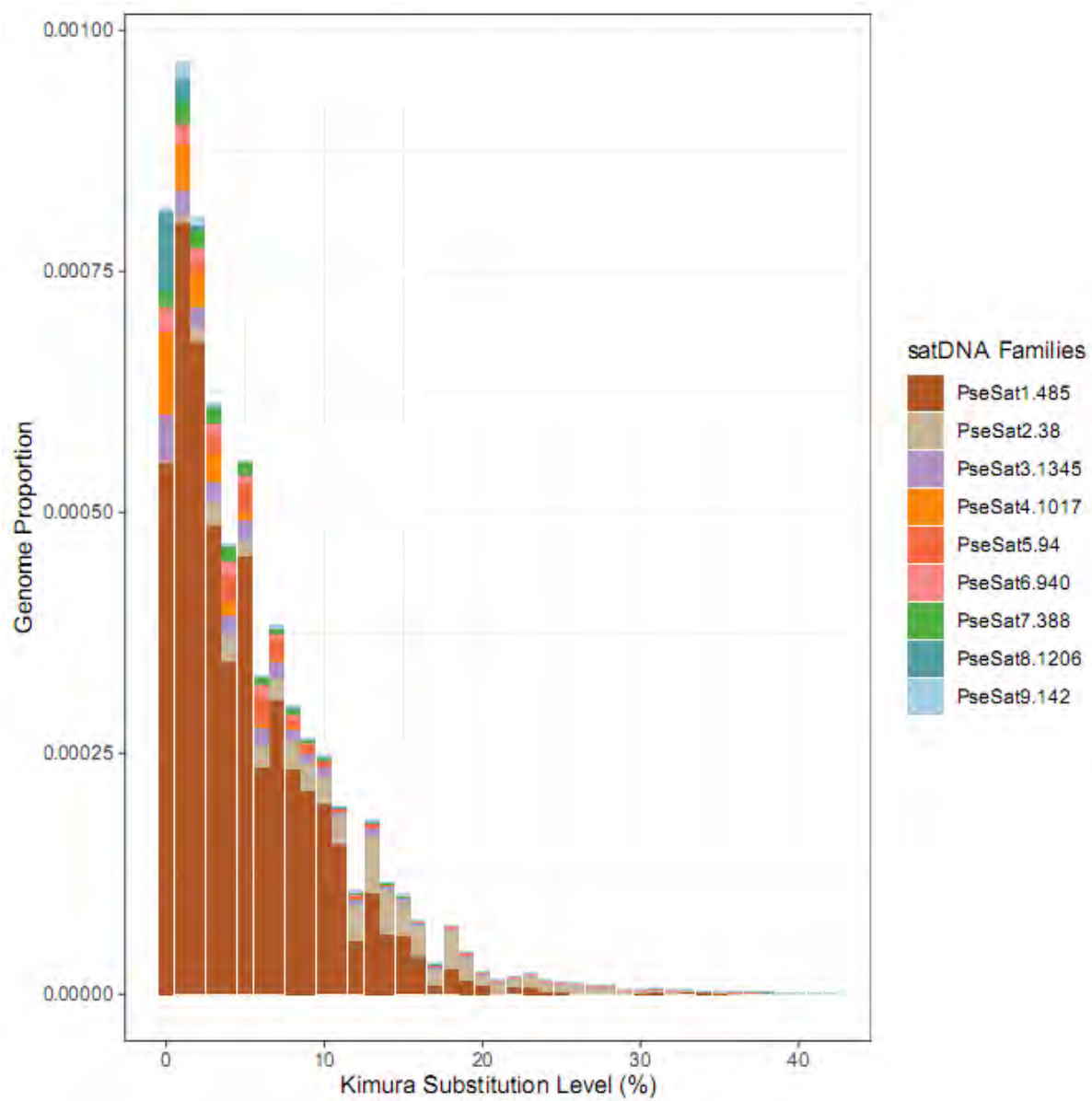


Figura S6. *RepeatLandscape* de *Polypterus senegalus*.

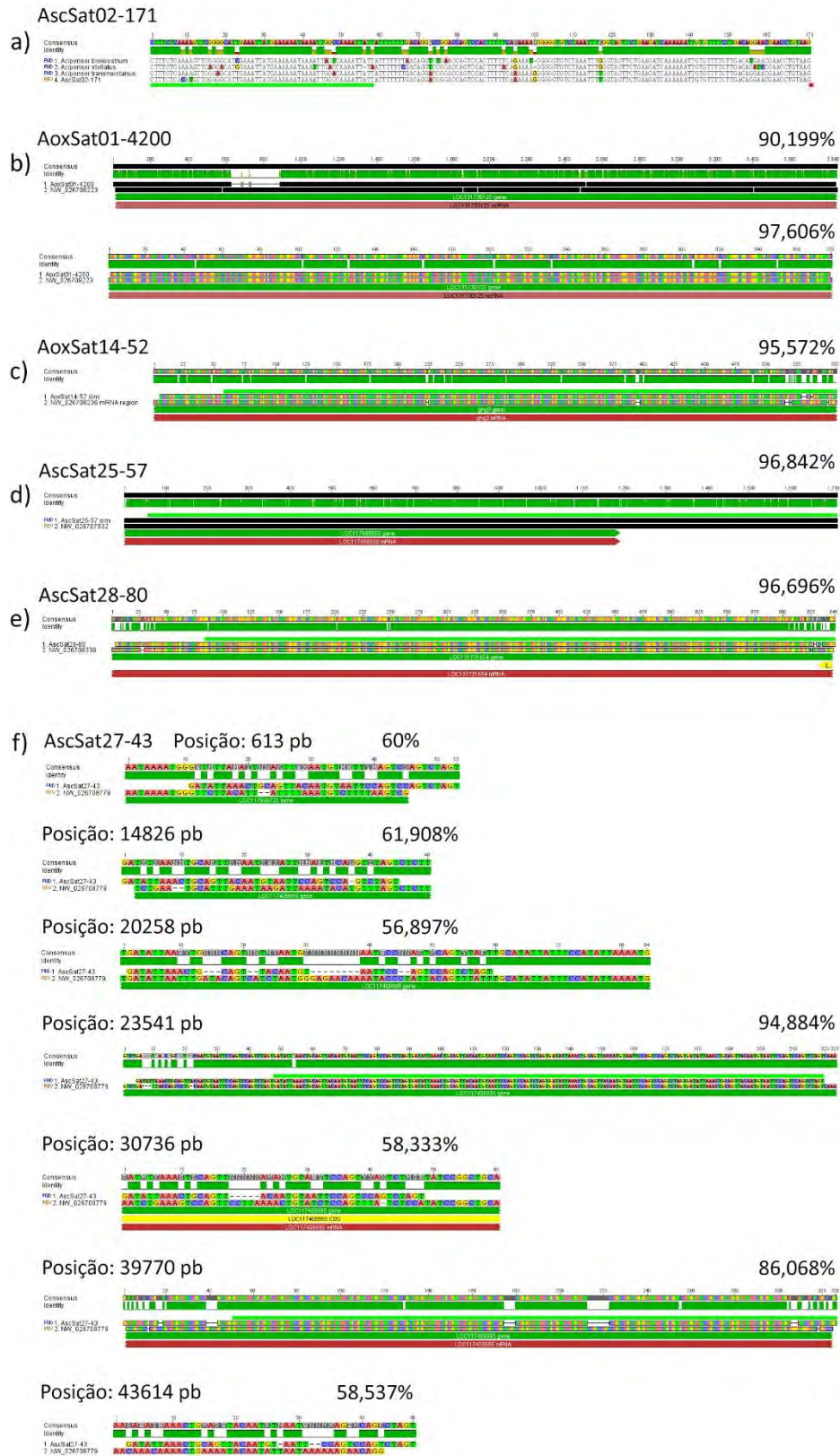


Figura S7. Alinhamentos dos resultados do BLAST. Em **a** foi realizado o alinhamento do AscSat02-171 com sequências de *HindIII* isoladas de outras espécies de esturjões, de **b – e** regiões que apresentaram repetições foram isoladas e alinhadas e em **f** diversas regiões apresentaram repetições, todas foram extraídas e alinhadas.