

UNIVERSIDADE FEDERAL DO PARANÁ  
VITOR CEDRAN PIRO  
MAURICIO TARABELLA GALVÃO

JFINISHER: FINALIZADOR DE MONTAGEM E EDITOR DE SEQUÊNCIAS  
BIOLÓGICAS

CURITIBA  
2012

VITOR CEDRAN PIRO  
MAURICIO TARABELLA GALVÃO

JFINISHER: FINALIZADOR DE MONTAGEM E EDITOR DE SEQUÊNCIAS  
BIOLÓGICAS

Trabalho de Conclusão de Curso apresentado ao  
Curso de Tecnologia em Análise e  
Desenvolvimento de Sistemas da Universidade  
Federal do Paraná como requisito à obtenção do  
título de Tecnólogo em Análise e Desenvolvimento  
de Sistemas

Orientador: Professor MSc. Dieval Guizelini  
Co-Orientador: Professor Dr. Roberto Tadeu Raittz

CURITIBA  
2012

## RESUMO

JFinisher é um software de alinhamento, edição e manipulação de sequências biológicas. Ele tem como objetivo auxiliar no processo de finalização de montagem de genomas. A partir de uma sequência de referência, o programa alinha contigs através do algoritmo de alinhamento local Smith-Waterman com métodos auxiliares, permitindo gerenciamento dos alinhamentos gerados. Ele possui interface gráfica para manipulação e visualização das ações, unindo funcionalidades que auxiliam na edição das sequências. Possui projetos internos gerenciáveis e capacidade de exportar resultados obtidos nos formatos padrão da área.

Palavras Chave: DNA, Genoma, montagem, bioinformática, alinhamento, Smith-Waterman

## ABSTRACT

JFinisher is software for alignment, editing and manipulation of biological sequences. It aims to assist in the finishing of genome assembly. Starting from a reference sequence, the program align contigs using Smith-Waterman local alignment algorithm with auxiliary methods, allowing management of the alignments generated. It has graphical interface for manipulation and visualization of the actions, uniting features that help in editing the sequences. It has internal projects manageable and ability to export results in the standard formats of the area.

Keywords: DNA, Genome, assembly, bioinformatics, alignment, Smith-Waterman

## LISTA DE ILUSTRAÇÕES

Figura 1 – Prototipação Incremental .....	18
Figura 2 – Tela principal jFinisher .....	24
Figura 3 – Menu “Arquivo” .....	25
Figura 4 – Seleção de arquivo referência.....	26
Figura 5 – Arquivo referência aberto .....	27
Figura 6 – Abrir contigs .....	28
Figura 7 – Tela principal com referência e contigs abertos .....	29
Figura 8 – Seleção de sequências .....	30
Figura 9 – Mapa com zoom.....	31
Figura 10 – Tela após rolagem horizontal .....	32
Figura 11 – Desativar opção “Travar barras”.....	33
Figura 12 – Barras de rolagem “destravadas” .....	34
Figura 13 – Tela de alinhamento.....	35
Figura 14 – Alinhando contigs .....	36
Figura 15 – Saída de alinhamento – Relatório por alinhamento.....	36
Figura 16 – Relatório geral de alinhamento.....	37
Figura 17 – Tela principal após alinhamento.....	38
Figura 18 – Múltiplos alinhamentos por contig .....	39
Figura 19 – Remover alinhamento .....	40
Figura 20 – Clique botão direito sobre sequência .....	41
Figura 21 – Remover contig .....	41
Figura 22 – Tela de entrada de dados .....	42
Figura 23 – Tela de navegação “Ir para” .....	42
Figura 24 – Reposicionamento de tela pela ocorrência encontrada.....	43
Figura 25 – Tela principal com projeto ativo.....	44

Figura 26 – Tela para unir sequências .....	44
Figura 27 – Menu exportar .....	45
Figura 28 – Diagrama de Caso de uso.....	51
Figura 29 – Diagrama de Classe.....	52
Figura 30 – Diagrama de Sequência – Abrir arquivo referência - Principal.....	88
Figura 31 – Diagrama de Sequência – Abrir arquivo referência – A1 .....	89
Figura 32 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência- contigs – Principal – A3 - A5 - A6 - A7 .....	90
Figura 33 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência- contigs – A1.....	91
Figura 34 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência- contigs – A2.....	92
Figura 35 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência- contigs – A4.....	93
Figura 36 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência- contigs – A8 – A9 .....	94
Figura 37 – Diagrama de Sequência – Remover sequência - Principal .....	95
Figura 38 – Diagrama de Sequência – Remover sequência – A1 .....	96
Figura 39 – Diagrama de Sequência – Navegar sequências-mapa - Principal .....	97
Figura 40 – Diagrama de Sequência – Localizar – Ir para posição – Principal – A1 ..	98
Figura 41 – Diagrama de Sequência – Editar sequência – Excluir - Principal.....	99
Figura 42 – Diagrama de Sequência – Editar sequência – Excluir – A1 .....	100
Figura 43 – Diagrama de Sequência – Editar sequência – Remover-Agrupar Ns – Principal .....	101
Figura 44 – Diagrama de Sequência – Editar sequência – Remover-Agrupar Ns – A1 .....	101
Figura 45 – Diagrama de Sequência – Editar sequência – Inserir – Principal.....	102
Figura 46 – Diagrama de Sequência – Editar sequência – Inserir – A1 .....	103
Figura 47 – Diagrama de Sequência – Alinhar sequência – Principal.....	104

Figura 48 – Diagrama de Sequência – Alinhar sequência – A1 .....	105
Figura 49 – Diagrama de Sequência – Alinhar sequência – A2 .....	106
Figura 50 – Diagrama de Sequência – Visualizar alinhamento - Principal .....	106
Figura 51 – Diagrama de Sequência – Exportar sequência – Principal .....	107
Figura 52 – Diagrama de Sequência – Exportar sequência – A1.....	108
Figura 53 – Diagrama de Sequência – Agrupar sequências – Principal – A1 – A2.	109
Figura 54 – Diagrama de Sequência – Agrupar sequências – A3.....	109
Figura 55 – Diagrama de Sequência – Agrupar sequências – A4.....	110
Figura 56 – Diagrama de Sequência – Novo projeto – Principal.....	111
Figura 57 – Diagrama de Sequência – Novo projeto – A1 .....	112
Figura 58 – Diagrama de Sequência – Abrir projeto – Principal.....	113
Figura 59 – Diagrama de Sequência – Abrir projeto – A1 .....	114
Figura 60 – Diagrama de Sequência – Salvar projeto - Principal.....	115
Figura 61 – Diagrama de Sequência – Salvar projeto – A1 .....	116
Figura 62 - WORK BREAKDOWN STRUCTURE .....	125
Figura 63 - GANTT.....	126

## **LISTA DE ABREVIATURAS E SIGLA**

DNA (ADN) – Ácido desoxirribonucleico

Fasta – Tipo de arquivo que armazena sequências

Gbk – Tipo de arquivo que armazena sequências com informações adicionais

GenBank – Base de dados pública de sequências genéticas mantida pela NCBI

GIT – Sistema de controle de versão

JAligner e JContigSort – Bibliotecas em Java para alinhamento local de sequências

JDK – Java Development Kit

JRE – Java Runtime Environment

NCBI – National Center for Biotechnology Information

RNA (ARN) – Ácido ribonucleico

UML - Unified Modeling Language

WBS - Work Breakdown Structure



## SUMÁRIO

RESUMO.....	II
ABSTRACT .....	III
LISTA DE ILUSTRAÇÕES .....	IV
LISTA DE ABREVIATURAS E SIGLA.....	VII
1 INTRODUÇÃO.....	11
1.1 INTRODUÇÃO .....	11
1.2 OBJETIVOS DO PROJETO.....	12
2 FUNDAMENTAÇÃO TEÓRICA .....	13
2.1 BIOINFORMÁTICA .....	13
2.2 DNA E SEQUENCIAMENTO .....	13
2.3 ALINHAMENTO DE SEQUÊNCIAS.....	14
2.4 ARQUIVOS FASTA E GENBANK.....	15
2.5 JAVA.....	15
2.6 UML.....	16
3 METODOLOGIA .....	17
3.1 MODELO DE PROCESSO DE ENGENHARIA DE SOFTWARE.....	17
3.2 PLANO DE ATIVIDADES.....	20
3.3 PLANO DE RISCOS .....	20
3.4 RESPONSABILIDADES.....	20
3.5 MATERIAIS.....	21
3.6 DESENVOLVIMENTO DO PROJETO .....	21
4 APRESENTAÇÃO DO SOFTWARE.....	23
4.1 INSTALAÇÃO .....	23
4.2 ABRIR ARQUIVOS .....	24
4.3 NAVEGAÇÃO .....	29

4.4	ALINHAMENTO .....	35
4.5	EDIÇÃO E LOCALIZAÇÃO .....	40
4.6	PROJETO .....	43
4.7	UNIR CONTIGS .....	44
4.8	EXPORTAR .....	45
5	CONSIDERAÇÕES FINAIS .....	46
	REFERÊNCIAS .....	49
	APÊNDICE .....	51
1	DIAGRAMA DE CASO DE USO.....	51
2	DIAGRAMA DE CLASSES .....	52
3	ESPECIFICAÇÃO DE CASOS DE USO .....	53
	UC01 – Abrir arquivo referência .....	53
	UC02 – Abrir-Adicionar arquivo(s) de sequência-contigs .....	56
	UC03 – Remover sequência.....	62
	UC04 – Navegar sequências-mapa.....	63
	UC05 – Localizar-Ir para posição .....	66
	UC06 – Editar sequência – Excluir .....	68
	UC07 – Editar sequência - Remover-Agrupar Ns .....	70
	UC08 – Editar sequência - Inserir .....	72
	UC09 – Alinhar sequência .....	73
	UC10 – Visualizar alinhamento.....	76
	UC11 – Exportar sequência.....	77
	UC12 – Agrupar sequências.....	80
	UC13 – Novo projeto .....	83
	UC14 – Abrir projeto .....	84
	UC15 – Salvar projeto.....	86
4	DIAGRAMAS DE SEQUÊNCIA.....	88

UC01 – Abrir arquivo referência .....	88
UC02 – Abrir-Adicionar arquivo(s) de sequência-contigs .....	90
UC03 – Remover sequência.....	95
UC04 – Navegar sequências-mapa.....	97
UC05 – Localizar-Ir para posição .....	98
UC06 – Editar sequência – Excluir .....	99
UC07 – Editar sequência - Remover-Agrupar Ns .....	101
UC08 – Editar sequência – Inserir .....	102
UC09 – Alinhar sequência .....	104
UC10 – Visualizar alinhamento.....	106
UC11 – Exportar sequência.....	107
UC12 – Agrupar sequências.....	109
UC13 – Novo projeto .....	111
UC14 – Abrir projeto .....	113
UC15 – Salvar projeto.....	115
<b>5 PGP – PLANO DE GERENCIAMENTO DE PROJETO .....</b>	<b>117</b>
1 INTRODUÇÃO .....	117
2 ESTIMATIVAS DO PROJETO .....	118
3 RISCOS DO PROJETO .....	121
4 CRONOGRAMA.....	125
5 RECURSOS DO PROJETO.....	127
6 ORGANIZAÇÃO DO PESSOAL.....	127
7 PLANO DE COMUNICAÇÃO.....	128
8 MECANISMOS DE RASTREAMENTO E CONTROLE.....	128

# 1 INTRODUÇÃO

## 1.1 INTRODUÇÃO

Neste trabalho, apresentamos os artefatos que documentam o desenvolvimento do aplicativo JFinisher - um produto que apresenta uma solução para uma demanda observada na área da bioinformática, onde os pesquisadores das áreas da bioquímica e informática precisam analisar e concluir os resultados dos processos de montagem de genomas. Ou seja, o software auxilia a finalização de montagem de genomas.

O trabalho de fechamento do genoma requer diversas atividades, uma das principais é o "alinhamento de sequências", que consiste em alinhar, analisar e manipular grandes quantidades de dados. Para isso, profissionais e estudantes utilizam diversos softwares, sendo grande parte deles abertos e/ou gratuitos, disponíveis na internet.

Para essa atividade existem diversas aplicações, baseadas em dois algoritmos, chamados alinhamento local e alinhamento global. O algoritmo do alinhamento local mais conhecido é o Smith-Waterman, e a implementação mais conhecida é o aplicativo BLAST fornecido pelo National Center for Biotechnology Information (NCBI). O algoritmo de alinhamento global mais conhecido é o Needleman-Wunsch, não utilizado nesse trabalho.

Apesar da grande quantidade e variedade de programas disponíveis, poucos abrangem todas as funcionalidades necessárias para a resolução dos problemas. Isso ocorre por alguns fatores: a maioria dos programas executam funções independentes, ou seja, não são unificados. É necessário então que, para executar uma tarefa ou chegar a um objetivo, vários programas sejam utilizados. Outro grande problema são os softwares sem interface gráfica, onde suas funções são executadas apenas por linha de comando com saídas em forma de texto. Isto acontece por: facilidade de desenvolvimento, necessidade de alto desempenho, utilização restrita, entre outros. A utilização de diversos programas também pode resultar na incompatibilidade de arquivos e resultados.

Os problemas acima citados e as limitações encontradas em alguns softwares demandam maior tempo de aprendizado e execução pelo usuário, resultando em um maior tempo de trabalho para solução dos problemas.

Estes fatores motivaram o desenvolvimento de um software unificado para alinhamentos e manipulação de múltiplas sequências, o JFinisher, que irá realizar suas funções através de uma interface gráfica de fácil manipulação. O programa está focado em análise, alinhamento, edição e finalização de montagem de genomas, a partir da entrada de um arquivo de referência e múltiplos arquivos de sequências (também chamadas de contigs).

## **1.2 OBJETIVOS DO PROJETO**

O projeto tem o objetivo de especificar o planejamento e o desenvolvimento de um software de manipulação de dados para a área de bioinformática, detalhando suas etapas, desde a fase de concepção até sua finalização.

O objetivo geral do JFinisher é alinhar e editar sequências biológicas em um ambiente gráfico, para auxiliar em algumas etapas na finalização de montagem de genomas recém-sequenciados.

O software possui os seguintes objetivos específicos:

- Abrir arquivos de formatos padrão da área (Fasta e GenBank), de sequências única e/ou múltiplas;
- Gerar uma interface para visualização, manipulação e navegação dos dados, com suporte a várias línguas;
- Alinhar e editar sequências;
- Construir sequência consenso a partir dos dados manipulados;
- Unificar sequências alinhadas;
- Abrir e salvar projeto interno;
- Exportar dados e sequências.

## **2 FUNDAMENTAÇÃO TEÓRICA**

### **2.1 BIOINFORMÁTICA**

A bioinformática é uma ciência que une várias linhas de conhecimentos: a engenharia de software, ciência da computação, estatística, matemática, biologia molecular e a bioquímica. Ela surgiu da crescente necessidade de análise de dados biológicos, proveniente dos avanços nos estudos do DNA nas últimas décadas e do aumento das máquinas de sequenciamento automático (identificação do DNA de organismos). Esta área se subdivide em diversas linhas de pesquisa, como: análise de sequências, alinhamento e busca de sequências, banco de dados de sequências, análise da expressão gênica, reconhecimento de genes, reconstrução de árvores filogenéticas, predição da estrutura da proteica, análise de mutações cancerígenas, genômica comparativa, modelagem de sistemas biológicos, sistemas de anotação de genomas, entre outras. Atualmente, considerada a era Pós-Genômica (PENA, 2005), a bioinformática faz mais do que apenas analisar dados, ela desempenha um papel importante na pesquisa médica e biológica.

### **2.2 DNA E SEQUENCIAMENTO**

A base de estudo da bioinformática é o DNA, uma sequência de pares de base (nucleotídeos) formados pelos quatro ácidos nucleicos: Adenina, Citosina, Guanina e Timina e que são representados, respectivamente, pelas letras A, C, G e T. A sequência de DNA, também denominada genoma, é encontrada nas células dos organismos vivos. Todas as atividades bioquímicas (multiplicação, reparação, produção energética, entre outras) dos organismos estão “descritas” e reguladas nos genomas, em sequências denominadas genes. Os genes descrevem a codificação das proteínas, enzimas.

O tamanho da sequência genoma, normalmente, é expressa em pares de bases (pb), em função que na genética e na biologia molecular um par de bases consiste em dois nucleotídeos opostos e complementares nas cadeias de ácido desoxirribonucleico (ADN) e ácido ribonucleico (ARN) que estão conectadas por pontes de hidrogênio.

A identificação das sequências de DNA são realizadas por máquinas sequenciadoras, que identificam os nucleotídeos contidos nos organismos e produzem arquivos com sequências de textos, formadas pelas letras A, C, G ou T. Este processo, apesar da crescente evolução, ainda é custoso e complicado. A sequência de DNA é muito extensa (genoma humano chega a aproximadamente 3 bilhões de pares de base) e difícil de ser sequenciada. Uma das maiores limitações das técnicas de sequenciamento encontra-se na limitação do número de bases que o processo consegue identificar por vez e na qualidade da sequência gerada. Após o processo de sequenciamento, utilizam-se softwares que realizam a montagem, tais como o Velvet (ZERBINO et al., 2008), que fornecem como saída do programa uma ou várias (mais de 100mil) sequências, com 2kb a 50kb de tamanho, dependendo dos dados de entrada, denominados contigs, e um número um pouco menor de sequências com falhas internas, denominados scaffolds.

## **2.3 ALINHAMENTO DE SEQUÊNCIAS**

O alinhamento de sequências busca utilizar métodos computacionais eficientes procurando semelhanças entre longas sequências de dados sendo, geralmente, uma menor (por exemplo um contig) e uma maior (genoma).

Uma das técnicas utilizadas para este fim é baseada em programação dinâmica. Ela realiza o alinhamento de sequências construindo uma matriz de comparação de prefixos das sequências que devem ser alinhadas. Esta técnica se divide em 2 métodos:

- Alinhamento global: alinha duas sequências por inteiro, utilizando toda sua extensão, proposta por Needleman e Wunsch (NEEDLEMAN et al., 1970).
- Alinhamento local: procura regiões de alta similaridade dentro da sequência, por segmentos parciais que alinhem entre si, proposta por Smith e Waterman (SMITH et al., 1981).

Estes alinhamentos possuem matrizes de score, produzidos pela somatória de pontos entre similaridades, diferenças, abertura de falhas (gap opening) e extensão de falhas (gap extension) da matriz de comparação. Com estas técnicas é possível prever qual o possível alinhamento entre sequências.

## 2.4 ARQUIVOS FASTA E GENBANK

Fasta e GenBank são tipos comuns de arquivos de texto formatados utilizados para armazenamento de sequências.

Arquivos FASTA são arquivos de base textual simples. Iniciam com uma linha de descrição e identificação iniciada pelo símbolo maior-que “>”, seguido de linhas de sequências. O cabeçalho tem alguns padrões para identificação do arquivo. A extensão do arquivo também serve como identificador. Por exemplo: “.fasta” são de sequências simples, enquanto “.ffn” identifica arquivos com múltiplas sequências, geralmente informando regiões codificantes da sequência.

Os arquivos GenBank (extensão “.gbk”) tem a capacidade de armazenar mais informações referentes a sequência. Possui duas partes: seção de anotação (iniciada pela notação LOCUS) e seção de sequência (iniciada pela notação ORIGIN). A seção de anotação contém: descrição, nome científico, taxonomia do organismo, entre outros. Pode ainda ter informações sobre características que identificam regiões codificantes e outros sítios de significância biológica (unidades de transcrição, sítios de mutação, etc.). Também tem a capacidade de armazenar a tradução de proteínas para regiões codificantes, além de referências bibliográficas.

## 2.5 JAVA

Java é uma linguagem de programação de alto nível e uma plataforma (ambiente de software que executa sobre uma plataforma de hardware). Segundo Peter Jandl Junior (2007), “O Java é uma linguagem de programação de propósito geral, concorrente, baseada em classes e orientada a objetos [...]”.

É uma linguagem altamente portátil. Isto se deve pelo fato de funcionar sobre uma plataforma, chamada JVM, tornando as aplicações escritas em java funcionais em todos os hardwares onde a JVM possui uma versão.

Possui três segmentos principais: JavaME (Java Micro Edition) voltado para dispositivos móveis; JavaEE (Java Enterprise Edition) voltada para desenvolvimento de aplicações corporativas complexas; JavaSE (Java Standard Edition) utilizada para aplicações de pequeno e médio porte, segmento que foi utilizado para produção deste projeto.



Java se destaca por ser orientada a objetos, com alta performance e suporte ao desenvolvimento multithread, com alta segurança e independente de plataforma.

## **2.6 UML**

Segundo Douglas Marcos Da Silva (2001), “A Unified Modeling Language (UML) é [...] uma linguagem padronizada para a modelagem de sistemas de software orientados a objetos, sendo adotada pela indústria de software como linguagem-padrão [...]”. É um conjunto de diagramas e especificações feitas para o planejamento e desenvolvimento de softwares.

Para este projeto foram utilizados diagramas de caso de uso, diagramas de classe e de sequência. A construção destes diagramas visa documentar o sistema que será desenvolvido, planejando suas funcionalidades e evitando erros futuros.

O diagrama de caso de uso, através de sua estrutura simples, permite organizar os requisitos do sistema, com uma visão geral de todo o problema. É utilizado em todo o projeto, desde sua concepção até os testes finais. O diagrama de classes representa a estrutura do sistema orientado a objetos, com suas classes e relacionamentos. Demonstra a estrutura do software. Já o diagrama de sequência permite a visualização sequencial das funcionalidades do sistema, baseados nas especificações de cada caso de uso. Isto permite, além de documentar, guiar o desenvolvimento do software conforme planejado.

### 3 METODOLOGIA

A metodologia utilizada para o desenvolvimento do projeto foi a prototipação incremental. Optou-se por este método pelo perfil do projeto. Por ser um trabalho de cunho científico e unificador, o projeto não possuía limites iniciais mensuráveis. Teríamos vários problemas e funcionalidades que poderiam ou não ser implementados. Dependendo do progresso do desenvolvimento e dos resultados obtidos teríamos novos requisitos.

#### 3.1 MODELO DE PROCESSO DE ENGENHARIA DE SOFTWARE

A metodologia de prototipação incremental adota o desenvolvimento por estágios. Cada estágio, quando finalizado, gera uma versão do software (Figura 1).

O processo incremental inicia com a concepção do software, onde as ideias são analisadas e discutidas para definir como será o software de maneira geral. Logo depois é feita a análise de requisitos, onde serão definidas as funcionalidades do sistema. Nesta etapa também é feito o projeto arquitetural, um modelo de alto nível de abstração que irá definir uma estrutura definida e clara para a solução do problema. Esta fase inicial ocorre de maneira cíclica, podendo ser revista diversas vezes antes do início do desenvolvimento.

Quando a primeira fase de projeto estiver em uma etapa avançada, já é possível iniciar o desenvolvimento do primeiro estágio, que irá abranger as principais funcionalidades do software.

“Quando um modelo incremental é usado, o primeiro incremento é frequentemente chamado de *núcleo do produto*. Isto é, os requisitos básicos são satisfeitos mas muitas características suplementares (algumas conhecidas, outras desconhecidas) deixam de ser elaboradas.”  
(PRESSMAN, 2006)

Neste estágio é comum revisar a primeira fase de projeto para garantir que a primeira versão estará funcional. Este ciclo se repete até a finalização do estágio. Com isso o sistema possui uma base funcional e com uma documentação bem especificada e está pronto para agregar as novas funcionalidades que virão nos próximos estágios.

Os estágios são divididos por funcionalidades completas e envolvem o planejamento, desenvolvimento, testes e entrega. Ao final de cada estágio é possível ter uma versão funcional do software. É necessário que as dependências entre os estágios sejam bem definidas para possam ser desenvolvidos conforme planejado.

O desenvolvimento de estágios segue até a finalização do projeto, ou vai sendo incrementado a cada novo requisito apresentado.

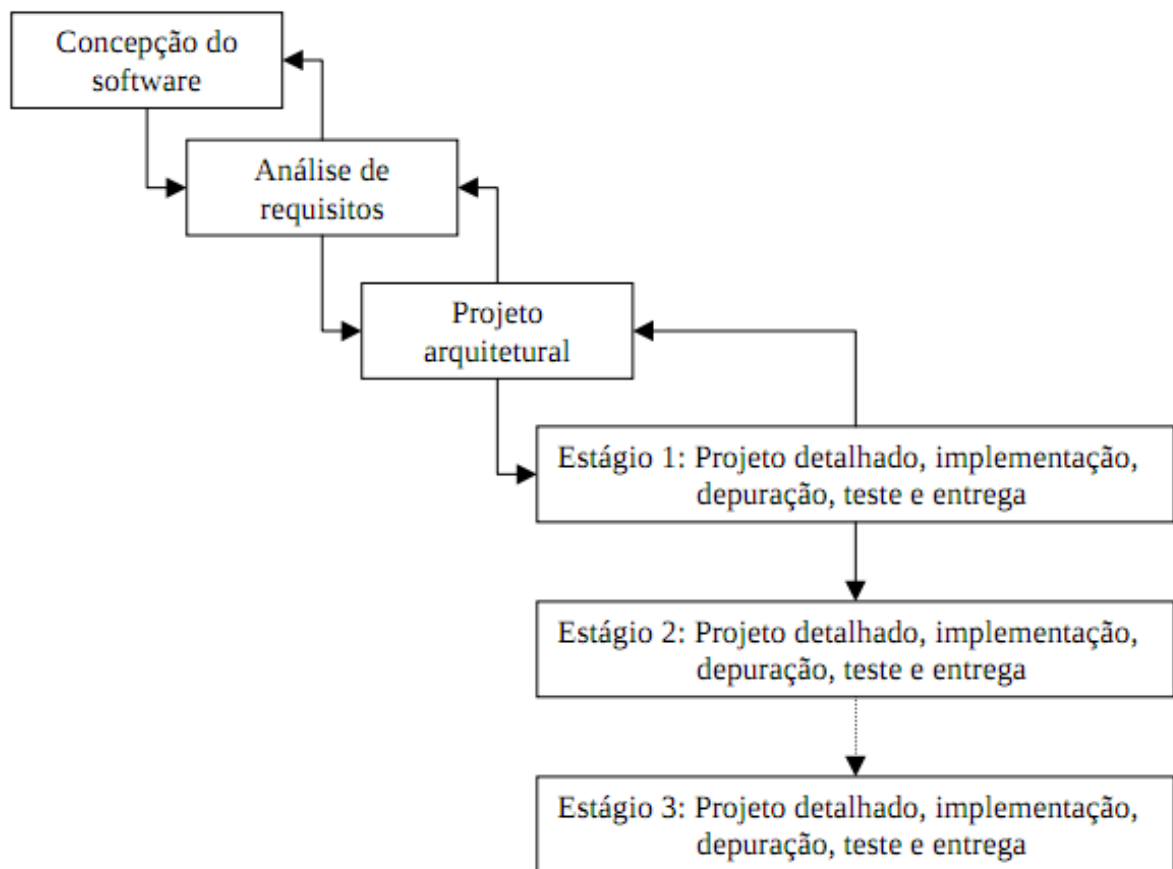


Figura 1 – Prototipação Incremental

Este modelo de prototipação foi a melhor opção para o desenvolvimento deste projeto. Isto ocorreu pelo perfil unificador do software planejado. Na fase de concepção existia a ideia geral do que seria desenvolvido, porém nem todas as funcionalidades que ele poderia agregar estavam definidas. Outro fator é a rapidez para desenvolvimento deste modelo, que permite alterações no projeto de forma gradual, reduzindo os riscos e dando flexibilidade, segmentado suas etapas.

Segue abaixo a divisão dos estágios no projeto do software JFinisher. A divisão está especificada de maneira detalhada no apêndice 5 PGP – PLANO DE GERENCIAMENTO DE PROJETO - 4.2 GRÁFICO DE GANTT.

- Estágio 1: Desenvolvimento da interface gráfica, navegação de sequências na tela, construção de menus gerais e estrutura multilíngual
- Estágio 2: Abertura de arquivos de sequência simples e múltiplos, seleção de sequências, construção do menu de opções
- Estágio 3: Alinhar sequências utilizando métodos JAligner e JContigSort, visualização de alinhamentos únicos
- Estágio 4: Remoção de sequências, construção do mapa de sequências, construção da sequência consenso
- Estágio 5: Desenvolvimento do múltiplo alinhamento por sequência
- Estágio 6: Gerenciamento dos projetos no software (novo, abrir, salvar)
- Estágio 7: Edição de sequências
- Estágio 8: Exportar sequências, navegação de posicionamento
- Estágio 9: Unificar sequências alinhadas
- Estágio 10: Revisão e finalização do software

O desenvolvimento da interface gráfica no estágio 1 e o alinhamento de sequências no estágio 3 foram as tarefas que mais demandaram esforços da equipe. São estágios que implementam funções chave para o sistema e tiveram que ser revisados diversas vezes para o correto funcionamento do programa. Consequentemente, estes estágios consumiram maior tempo de desenvolvimento.

O primeiro e segundo estágio forneceram uma base para a aplicação. Nesta etapa foram desenvolvidos os métodos para manipulação e visualização das sequências. Desenvolvemos um algoritmo para imprimir as sequências em tela, tarefa que demandou alto desempenho da aplicação. A biblioteca BioInfoCommons foi utilizada no processo de abertura dos arquivos.

No estágio 3 foram utilizadas as bibliotecas JContigSort e JAligner. Aplicamos o método de alinhamento local do JAligner em uma região restrita (dobro do comprimento da sequência selecionada), localizada nos extremos da mediana retornada pelo alinhamento do JContigSort.

No estágio 4 e 5 foram utilizadas técnicas de layouts proporcionais no Java que permitiram o desenvolvimento do mapa gráfico. A biblioteca XStream foi a base para o estágio 6, onde criamos a estrutura de projetos internos no sistema. Ela facilitou o armazenamento dos objetos, transformado em xml (e recuperando os arquivos previamente salvos). Nesta etapa utilizamos verificação de chave (MD5) para confirmar a integridade dos arquivos.

Os estágios 7 ao 10 foram relativamente menos complexos. Quase toda a estrutura do software estava completa, facilitando o acesso aos componentes internos. Estas últimas etapas e a finalização ocorreram em menos tempo que as demais.

### **3.2 PLANO DE ATIVIDADES**

As atividades foram divididas considerando a média de trabalho de cada participante distribuída dentro do tempo corrido do projeto. Isto ocorreu pois o tempo de dedicação ao projeto não era fixo no dia dos participantes. Assim, efetuou-se uma média de tempo de dedicação semanal por participante, que foi distribuída pelo calendário.

As etapas e tarefas desenvolvidas no projeto estão especificadas nos capítulos 4.1 WORK BREAKDOWN STRUCTURE– DIVISÃO DE TRABALHO NO PROJETO e 4.2 GRÁFICO DE GANTT do apêndice 5 PGP – PLANO DE GERENCIAMENTO DE PROJETO.

### **3.3 PLANO DE RISCOS**

O plano de riscos está descrito no apêndice no apêndice 5 PGP – PLANO DE GERENCIAMENTO DE PROJETO - 3 RISCOS DO PROJETO.

### **3.4 RESPONSABILIDADES**

As responsabilidades do projeto estão especificadas no apêndice 5 PGP – PLANO DE GERENCIAMENTO DE PROJETO - 6 ORGANIZAÇÃO DO PESSOAL.

### **3.5 MATERIAIS**

As características de hardware e software utilizados para o desenvolvimento do projeto estão especificadas no apêndice 5 PGP – PLANO DE GERENCIAMENTO DE PROJETO - 5 RECURSOS DO PROJETO.

### **3.6 DESENVOLVIMENTO DO PROJETO**

O início do projeto foi marcado por algumas dificuldades iniciais: a escolha do escopo e entendimento do tema. Por se tratar de um software de bioinformática, algum conhecimento de biologia molecular era necessário para o desenvolvimento correto do sistema. Através de estudos e análises de softwares semelhantes conseguimos lidar com esta dificuldade.

Após a decisão de se desenvolver um editor de sequências, ocorreram algumas dificuldades na escolha das funcionalidades, tendo em vista o limitado conhecimento na área. Não tínhamos ideia do que realmente poderia ser feito no tempo disponível. Iniciamos o projeto com algumas funcionalidades básicas, mas sempre tivemos uma grande lista de outras funções que poderiam ser agregadas. Ao decorrer do desenvolvimento conseguimos visualizar o que poderia e não ser feito, dentro do tempo previsto.

A escolha da linguagem também foi um fator importante no projeto, pois o software iria precisar de alto desempenho (trabalhar com sequências de milhões de caracteres), além de necessitar de algumas bibliotecas específicas para realizar alinhamentos e manipulação de arquivos. A solução mais adequada para o software, segundo nossa análise, foi o Java, onde já tínhamos conhecimento e atendia as necessidades solicitadas.

Como a equipe possui apenas 2 integrantes, optou-se a divisão entre desenvolvimento e planejamento. Um integrante faria toda a parte de desenvolvimento, enquanto o outro gerenciaria o projeto. Desta forma conseguimos ter um andamento mais linear e sem atrasos, com focos bem definidos. Porém os integrantes interagem constantemente em suas funções.

Tivemos um risco inicial em relação as datas para o desenvolvimento do projeto. Tínhamos a definição de conclusão para final de 2011, e estávamos realizando todo o planejamento para esta data, apesar de ser um tempo muito curto,

considerando que o projeto se iniciou em setembro do mesmo ano. Porém, ocorreu uma mudança desta data para final de fevereiro de 2012, acrescentando mais 3 meses para o desenvolvimento e finalização do software. Isto nos permitiu ampliar o escopo do desenvolvimento, além de conseguir atingir outro nível de maturidade com o projeto.

As reuniões com o orientador foram semanais, auxiliando no andamento do projeto. Utilizamos o GIT como ferramenta para gerenciamento de versões durante o desenvolvimento do software.

Por utilizarmos o método de prototipação incremental, gerenciamos o tempo de acordo com os estágios de desenvolvimento. Conseguimos manter as datas de entrega de cada estágio, apesar dos integrantes não possuírem disponibilidade integral para o projeto. Isso rendeu alguns atrasos em tarefas específicas, mas os estágios foram completados em tempo sem maiores dificuldades.

## 4 APRESENTAÇÃO DO SOFTWARE

### 4.1 INSTALAÇÃO

Para executar o software é necessário possuir instalada o JRE versão 6 ou superior. O JFinisher não necessita de instalação. O software é compactado no formato JAR. Para executá-lo abra o arquivo `jFinisher.jar` com o Java ou execute o comando: `java -jar jFinisher.jar` (comando deve ser executado no diretório em que o arquivo se encontra. Neste diretório deve haver a pasta “lib” com as bibliotecas necessárias que vem junto do software).

No CD de instalação que acompanha este documento o arquivo `jFinisher.jar` está no diretório `jFinisher`. Pode ser executado diretamente ou através dos arquivos `.bat` (para Windows) ou `.sh` (Linux).



## 4.2 ABRIR ARQUIVOS

Ao abrir o arquivo jFinisher.jar, será apresentada a tela principal do sistema, (Figura 2).

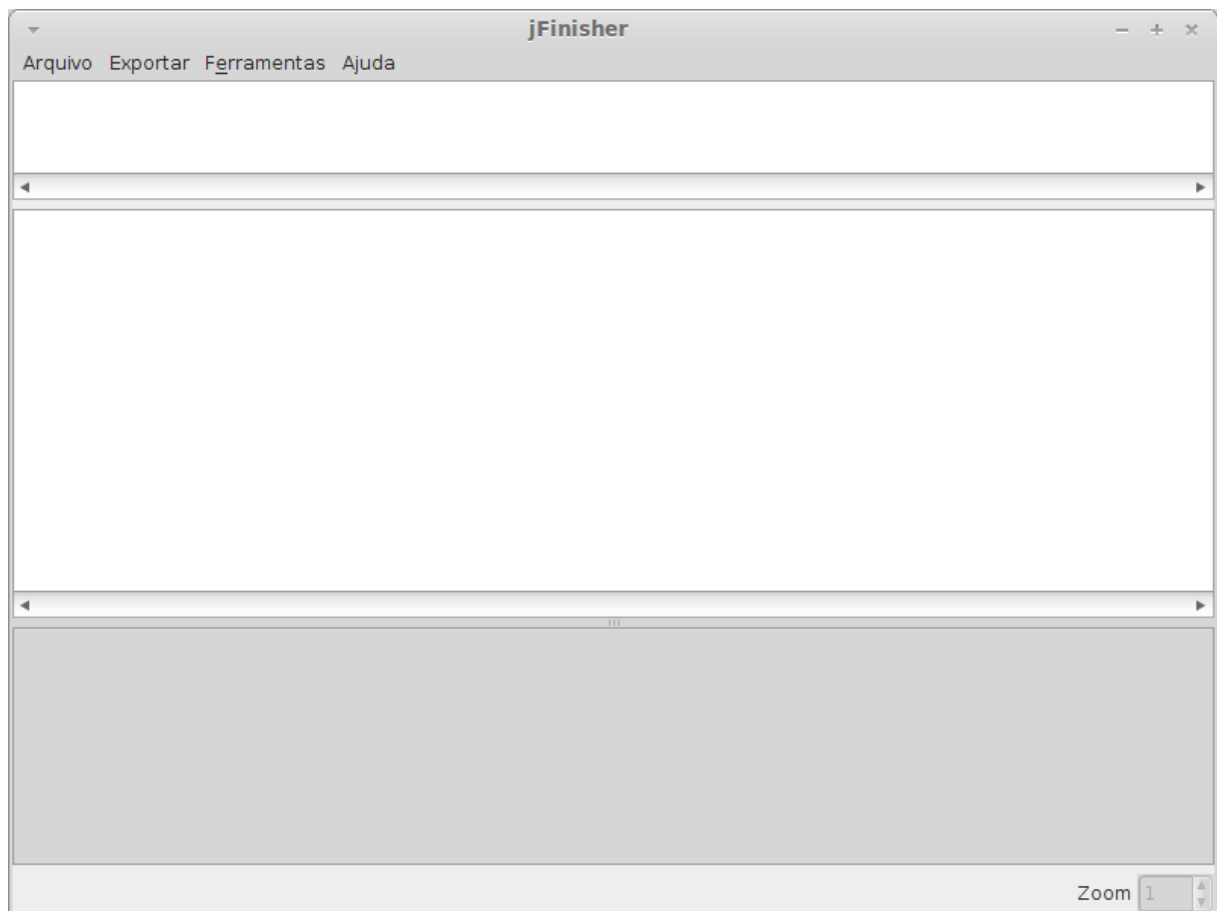


Figura 2 – Tela principal jFinisher

Através do menu superior do sistema é possível executar as suas funções. Pelo menu “Arquivo” é possível abrir sequência de referência ou abrir algum projeto previamente salvo (Figura 3).

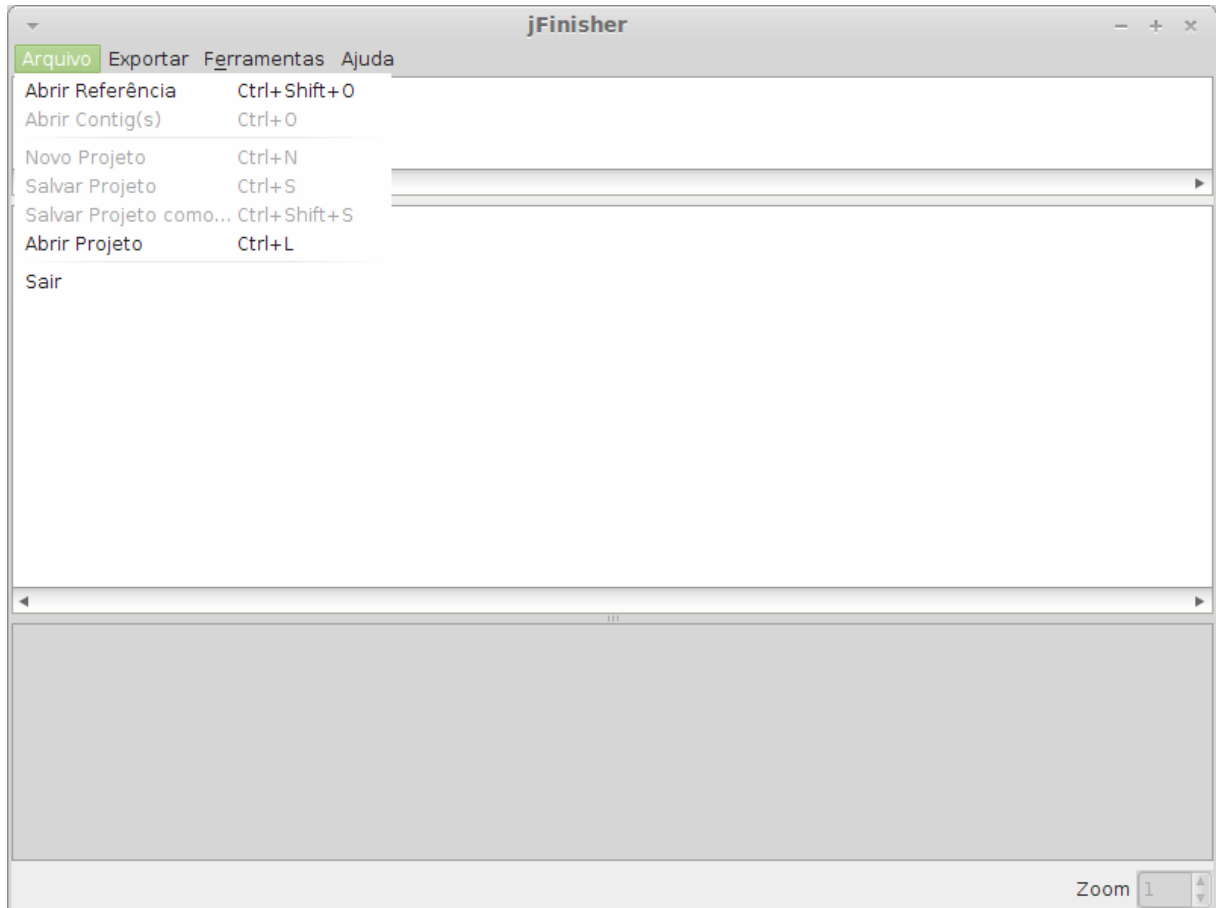


Figura 3 – Menu “Arquivo”

Ao selecionar “Abrir Referência” o usuário deve escolher o arquivo desejado em um formato permitido (Figura 4).

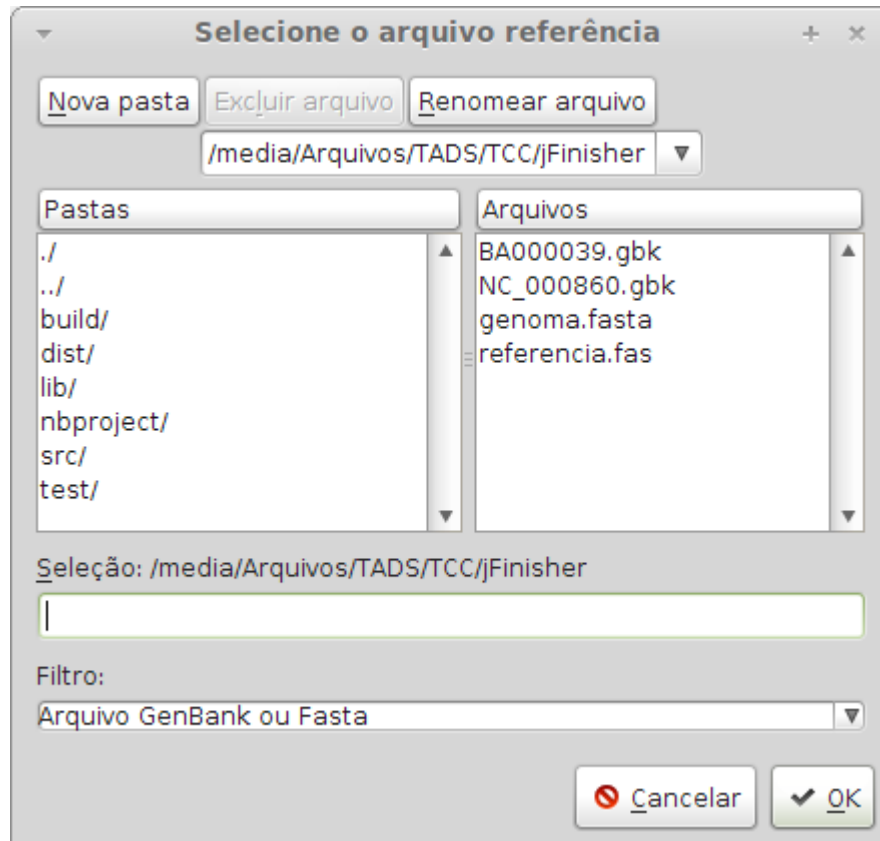


Figura 4 – Seleção de arquivo referência

Ao selecionar um arquivo de um formato permitido, sistema irá abrir na tela os dados (Figura 5).

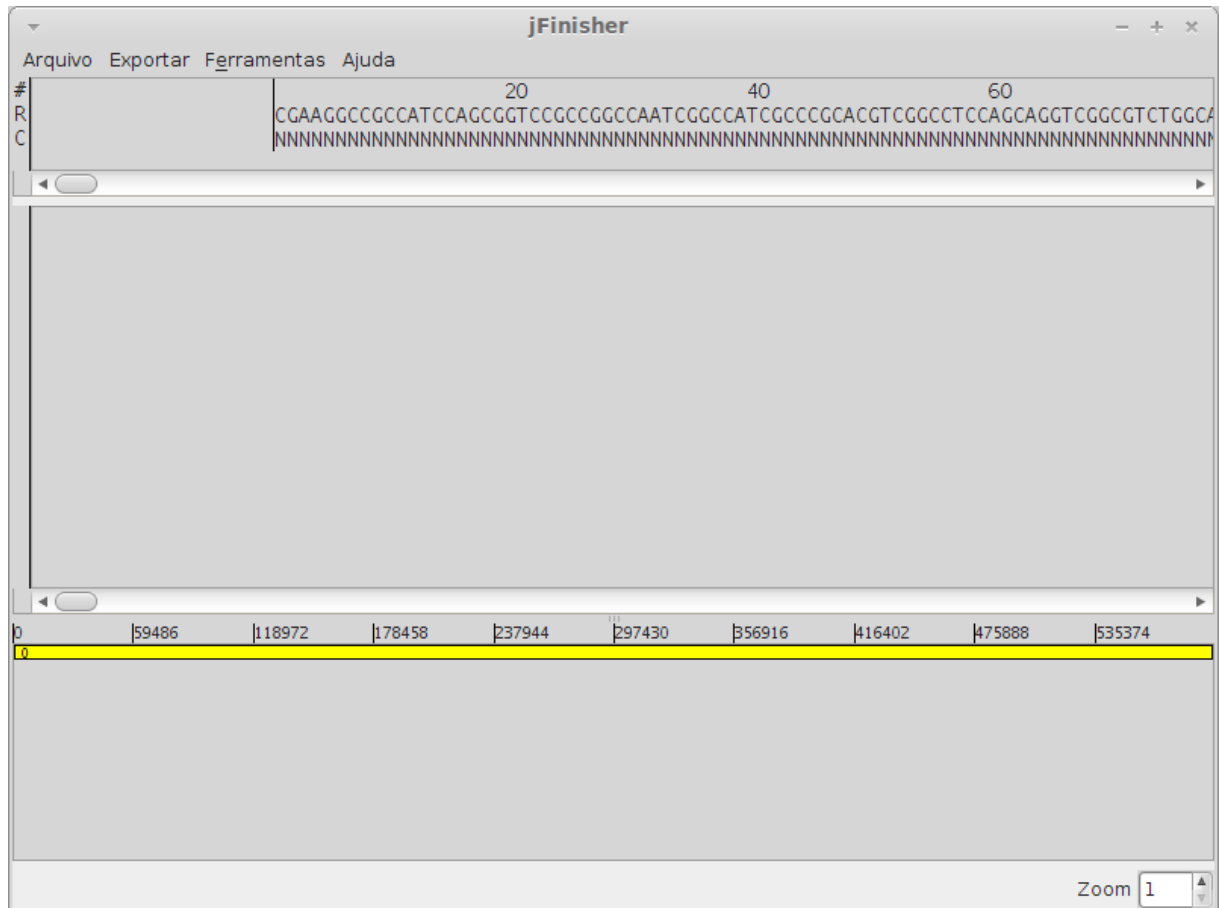


Figura 5 – Arquivo referência aberto

Para abrir contigs é necessário selecionar a opção “Abrir contig(s)” no menu “Arquivo” e selecionar um ou mais arquivos de formato permitido na tela de seleção (Figura 6).

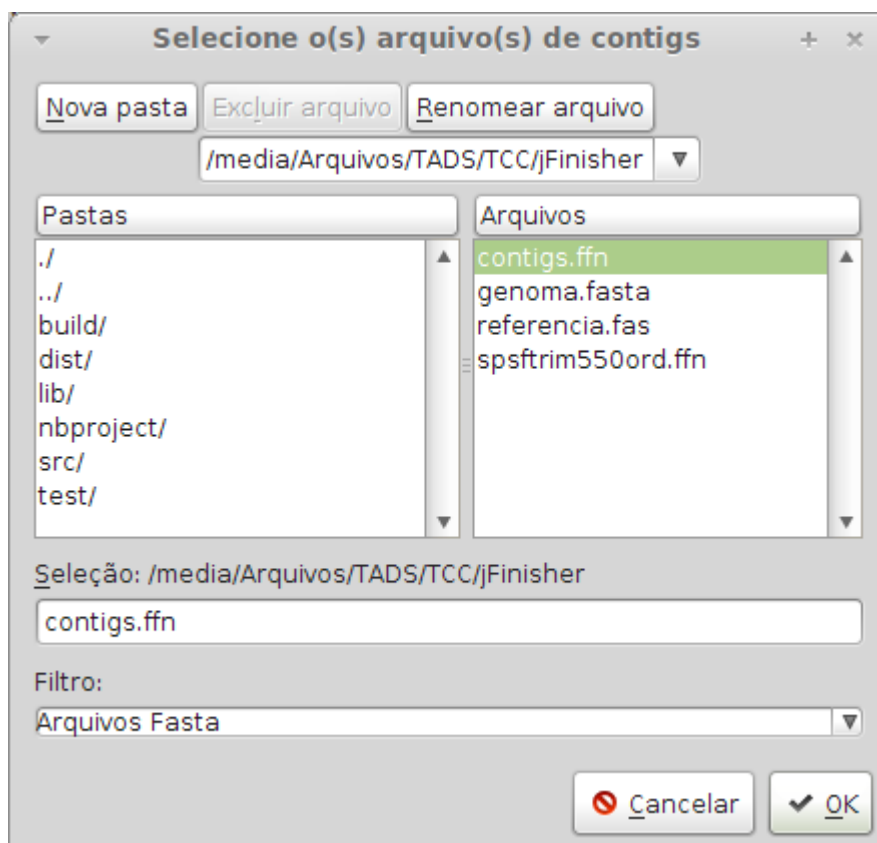


Figura 6 – Abrir contigs

Sistema abre e apresenta na tela os dados do(s) arquivo(s) de contig(s) (Figura 7).

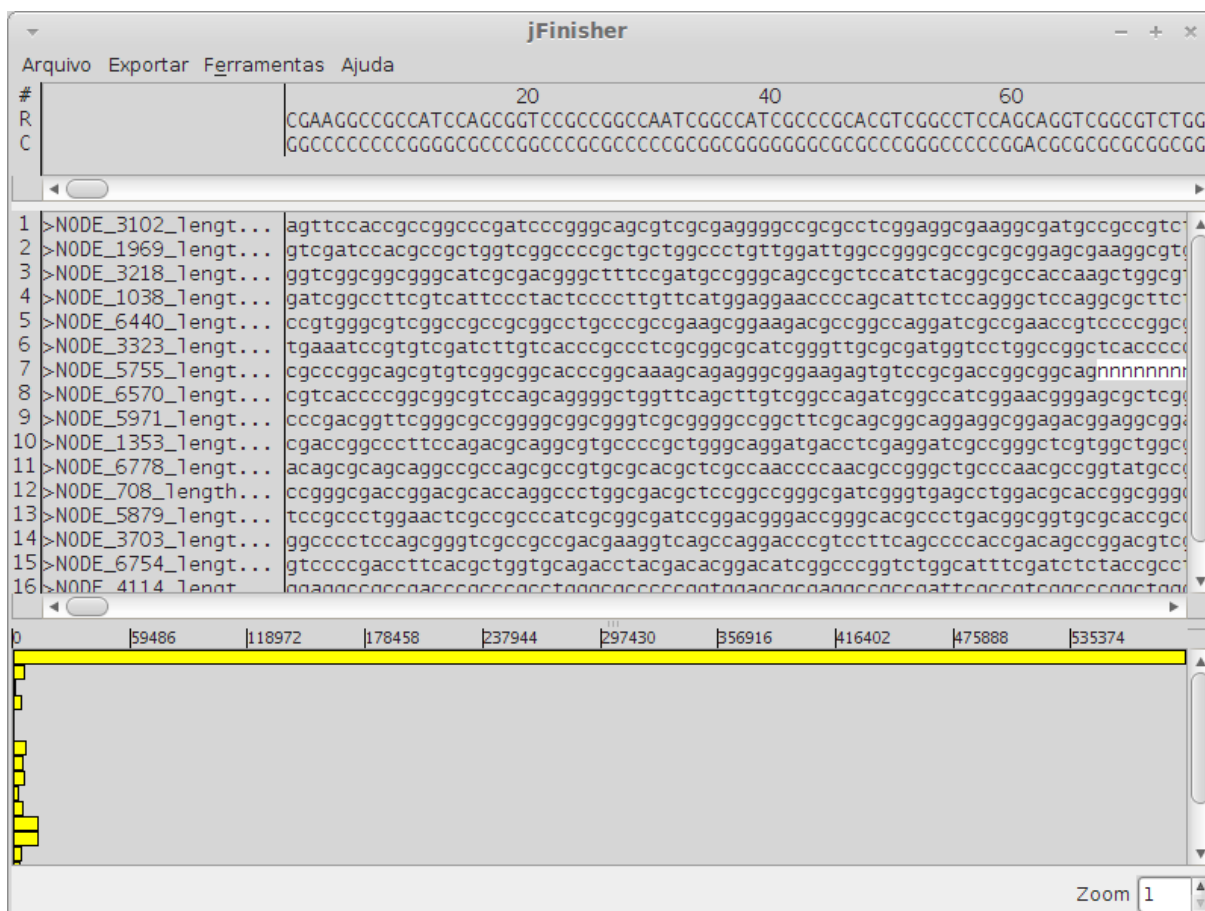


Figura 7 – Tela principal com referência e contigs abertos

Após abrir os dois tipo de arquivos (referência e contigs) já é possível realizar todas as funções do software. É possível também adicionar mais arquivos de contigs caso necessário.

### 4.3 NAVEGAÇÃO

A tela principal do software apresenta 3 divisões:

- Superior: contém contador, sequência de referência e sequência consenso
- Centro: todos os contigs com sua descrição e identificador.
- Inferior: mapa gráfico dos dados abertos no programa com contador e controlador de zoom. A linha maior que ocupa todo o comprimento da

tela representa a sequência referência. As outras representam os contigs com tamanho relativo a sequência principal.

Ao passar o cursor sobre as sequências será apresentado a posição do caractere dentro do arquivo (após alinhado esta informação aparece juntamente da posição de alinhamento). Ao passar o cursor e clicar (ou arrastar) é possível selecionar partes específicas de cada sequência (Figura 8).

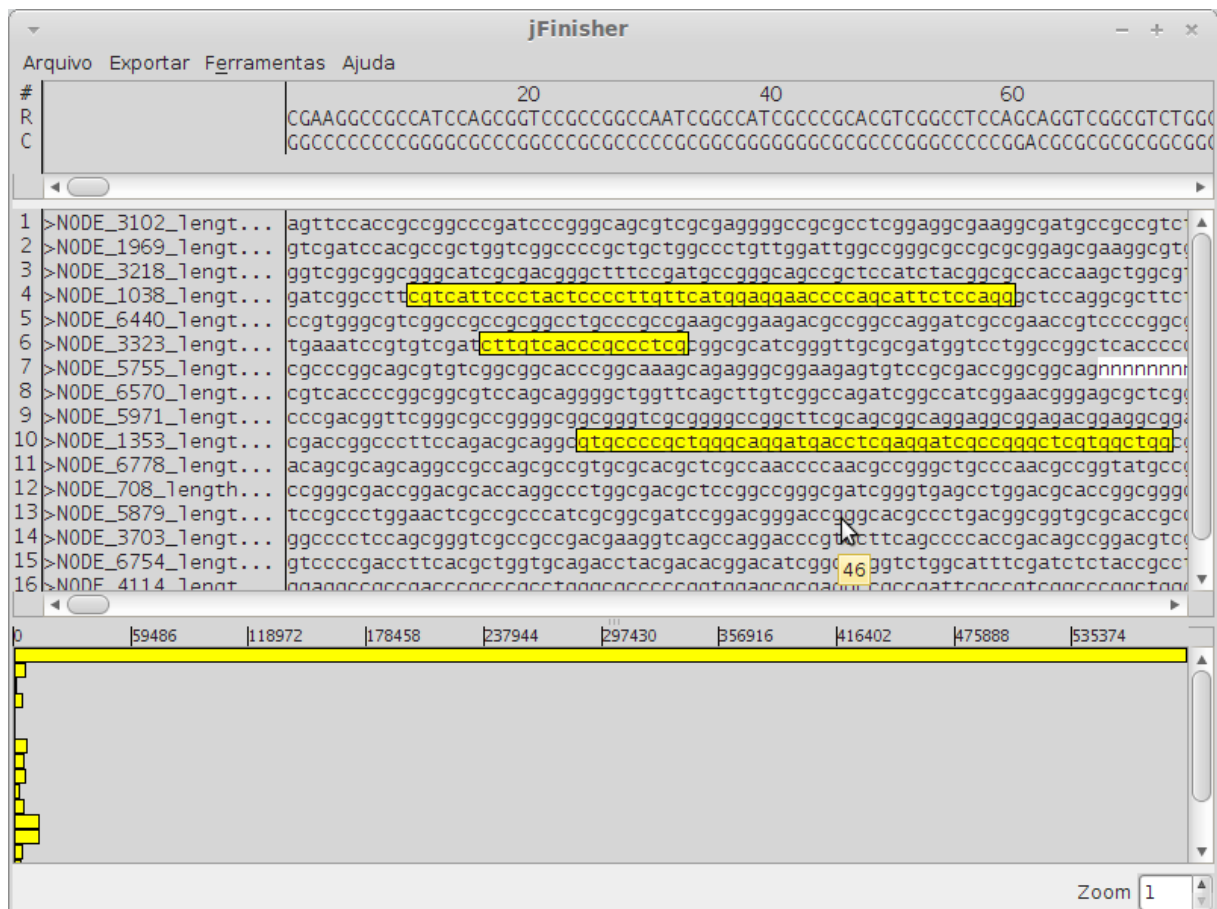


Figura 8 – Seleção de sequências

Os caracteres “N” são destacados em branco.

Alterando o valor do campo “Zoom” é possível ter uma maior definição do mapa. O contador do mapa é atualizado a cada alteração no valor do zoom para manter a proporção entre o tamanho das sequências (Figura 9).

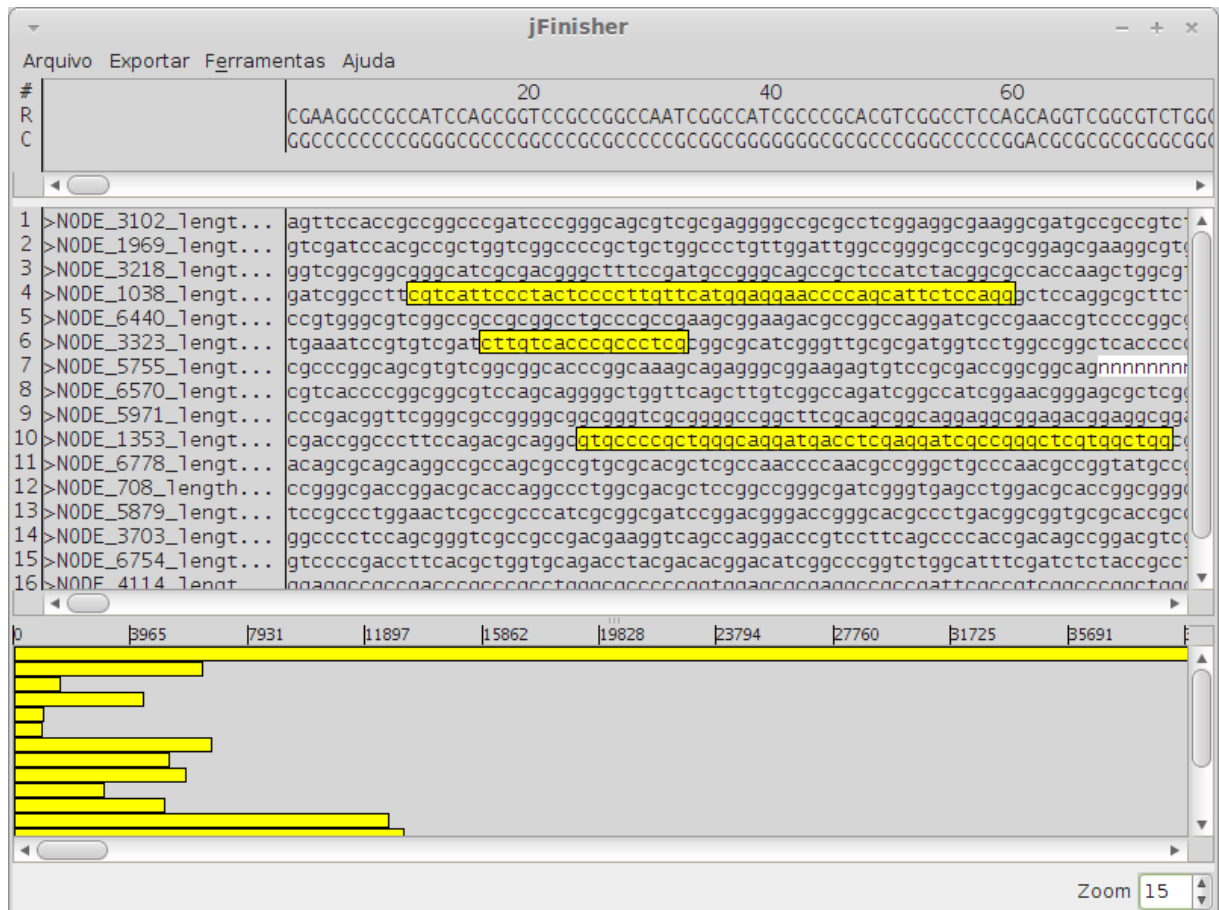


Figura 9 – Mapa com zoom



Ao efetuar a rolagem horizontal em qualquer barra, o sistema automaticamente posiciona todas as telas, fazendo com que a posição das sequências mantenham-se alinhadas (Figura 10).

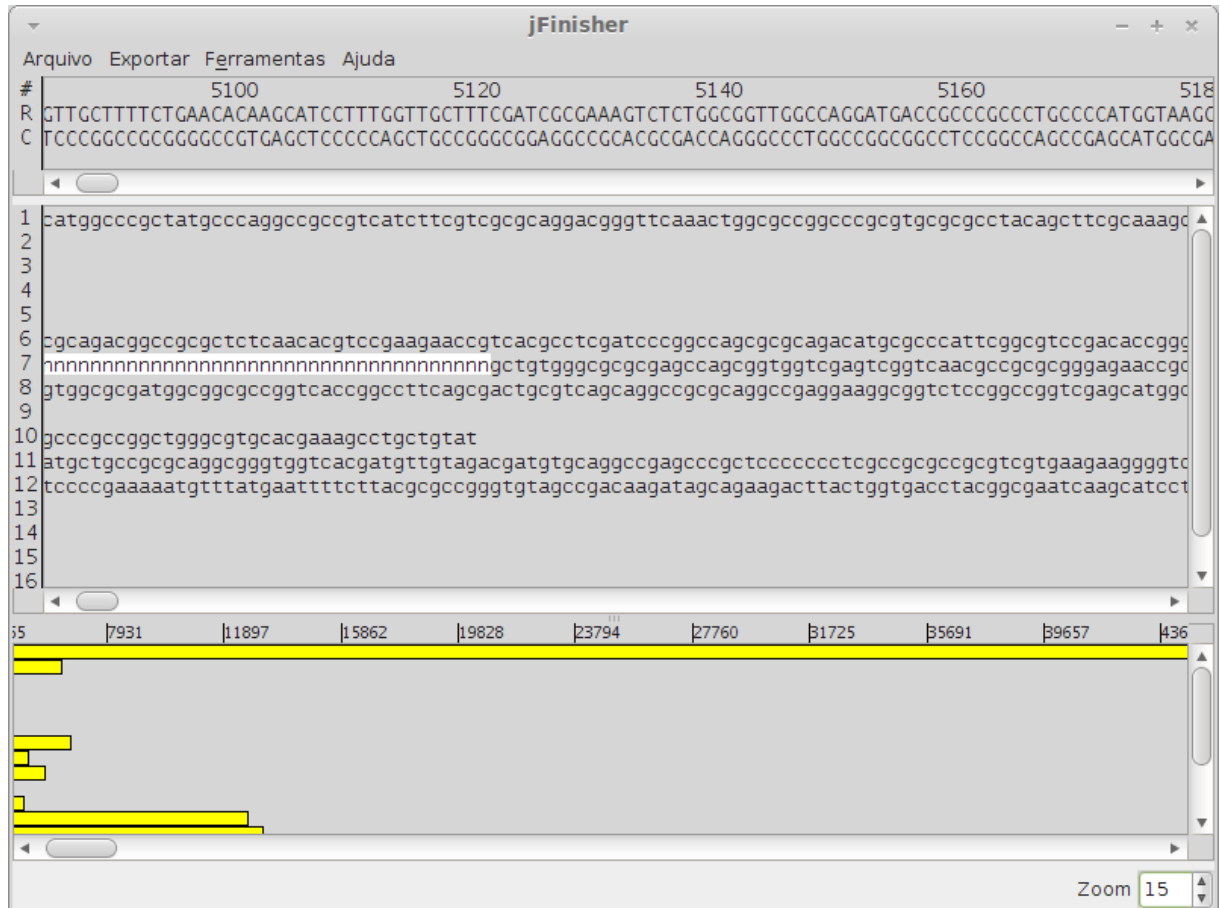


Figura 10 – Tela após rolagem horizontal

Desmarcando a opção “Travar barras” no menu “Ferramentas” (Figura 11) é possível navegar por qualquer área do sistema independente da outra (Figura 12)

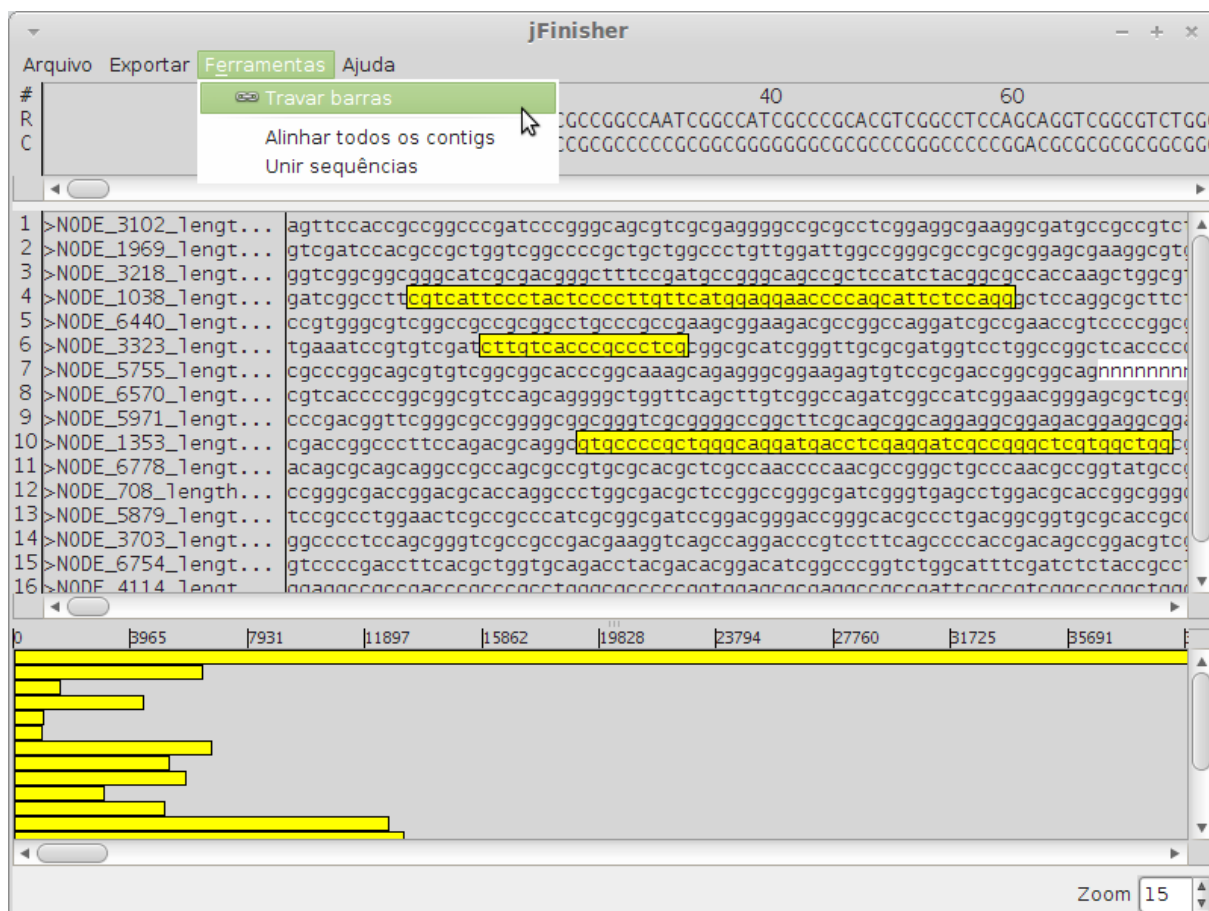


Figura 11 – Desativar opção “Travar barras”

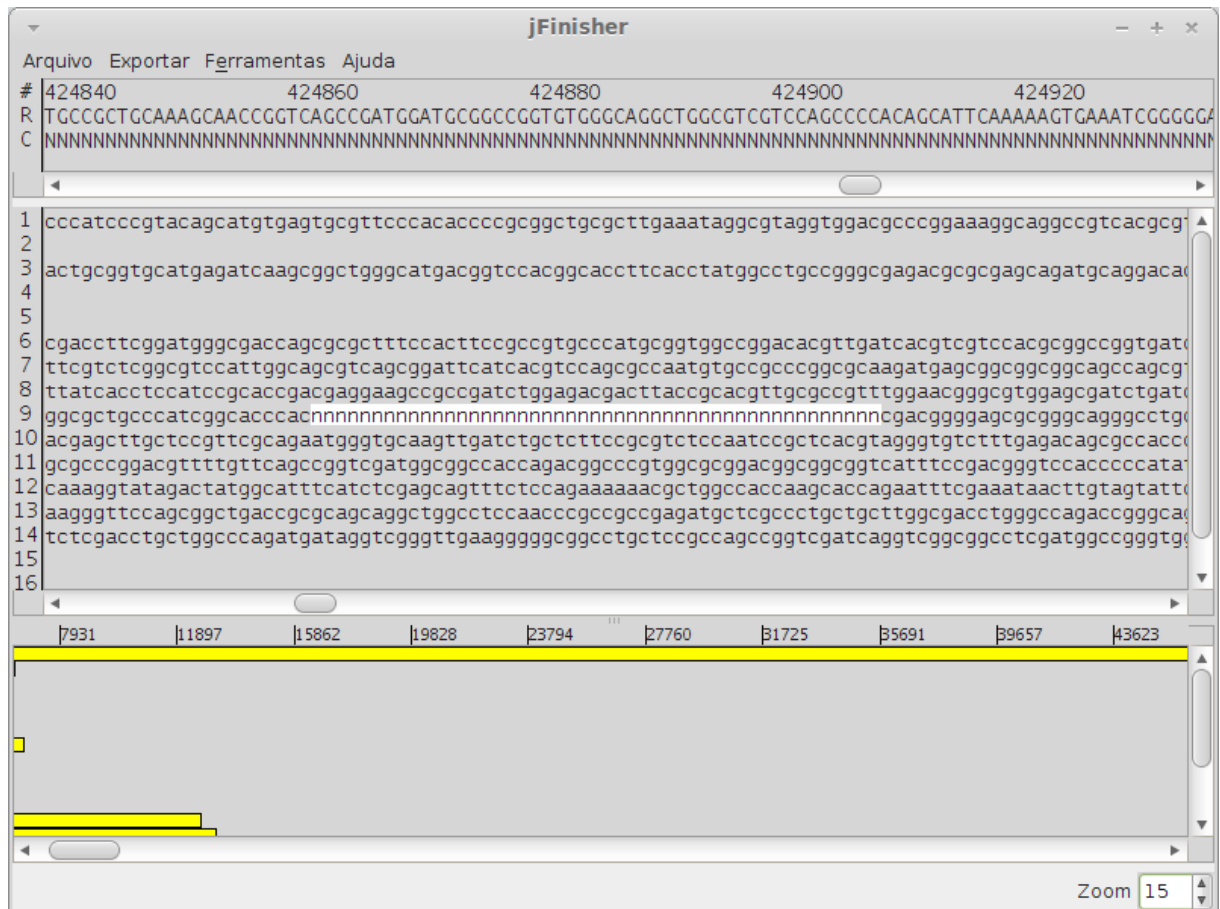


Figura 12 – Barras de rolagem “destravadas”

#### 4.4 ALINHAMENTO

O alinhamento pode ser efetuado individualmente por contig (clique com botão direito sobre sequência, opção “Novo Alinhamento”) ou realizada em todos os contigs (menu “Ferramentas” opção “Alinhar todos os contigs”). Realizando uma destas opções o sistema irá apresentar a tela de alinhamento (Figura 13).

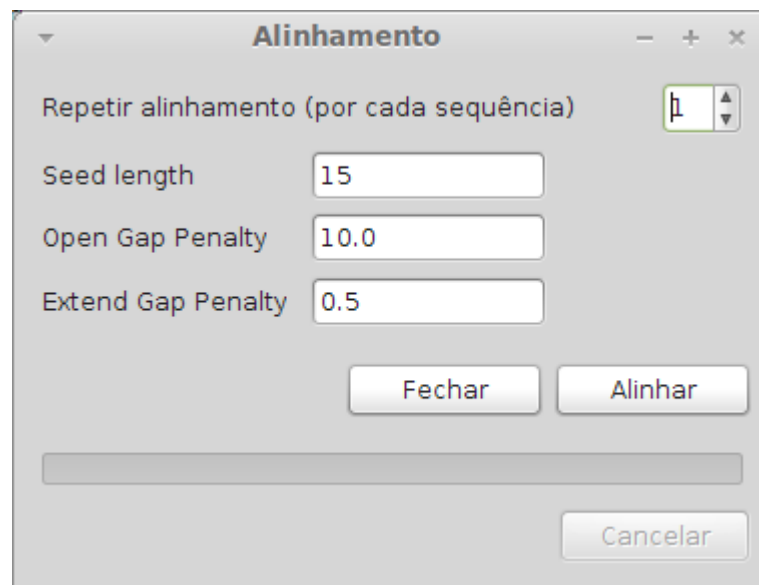


Figura 13 – Tela de alinhamento

Nesta tela é possível informar quantos alinhamento serão feitos (por contig) e os parâmetros para os métodos de alinhamento. Ao clicar no botão “Alinhar” o sistema alinha o(s) contig(s) selecionado(s) (Figura 14) e apresenta o relatório ao final do processo (Figura 15).

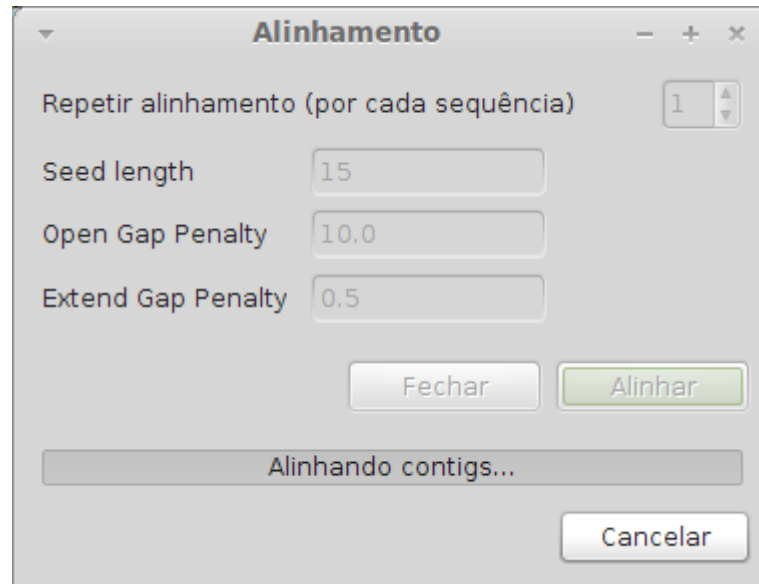


Figura 14 – Alinhando contigs

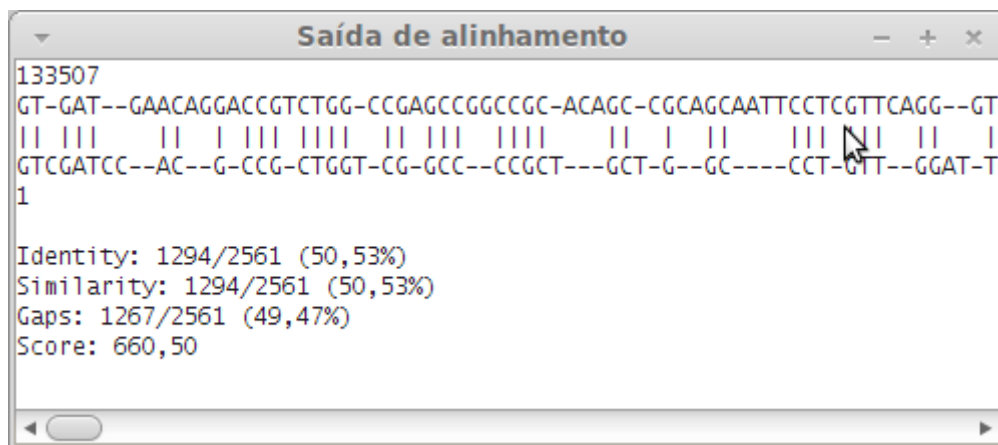


Figura 15 – Saída de alinhamento – Relatório por alinhamento

Ao realizar alinhamentos múltiplos o sistema gera um outro relatório (Figura 16), mostrando o status dos alinhamentos realizados, identificando cada um pelo número do contig.

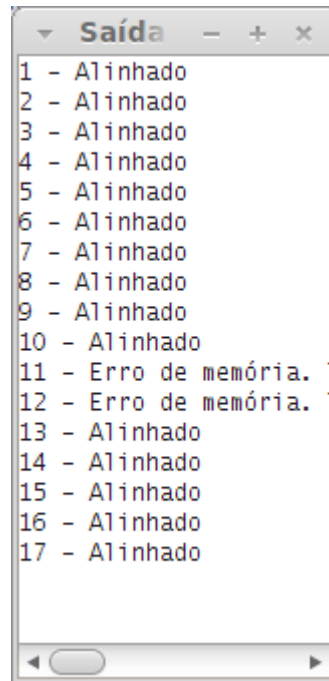


Figura 16 – Relatório geral de alinhamento



Ao realizar mais de um alinhamento por contig, o sistema irá distribuir os alinhamentos pelo mapa, tendo sempre apenas uma deles selecionado e ativo. Também é possível acessar o alinhamento clicando com o botão direito sobre a sequência (Figura 18).

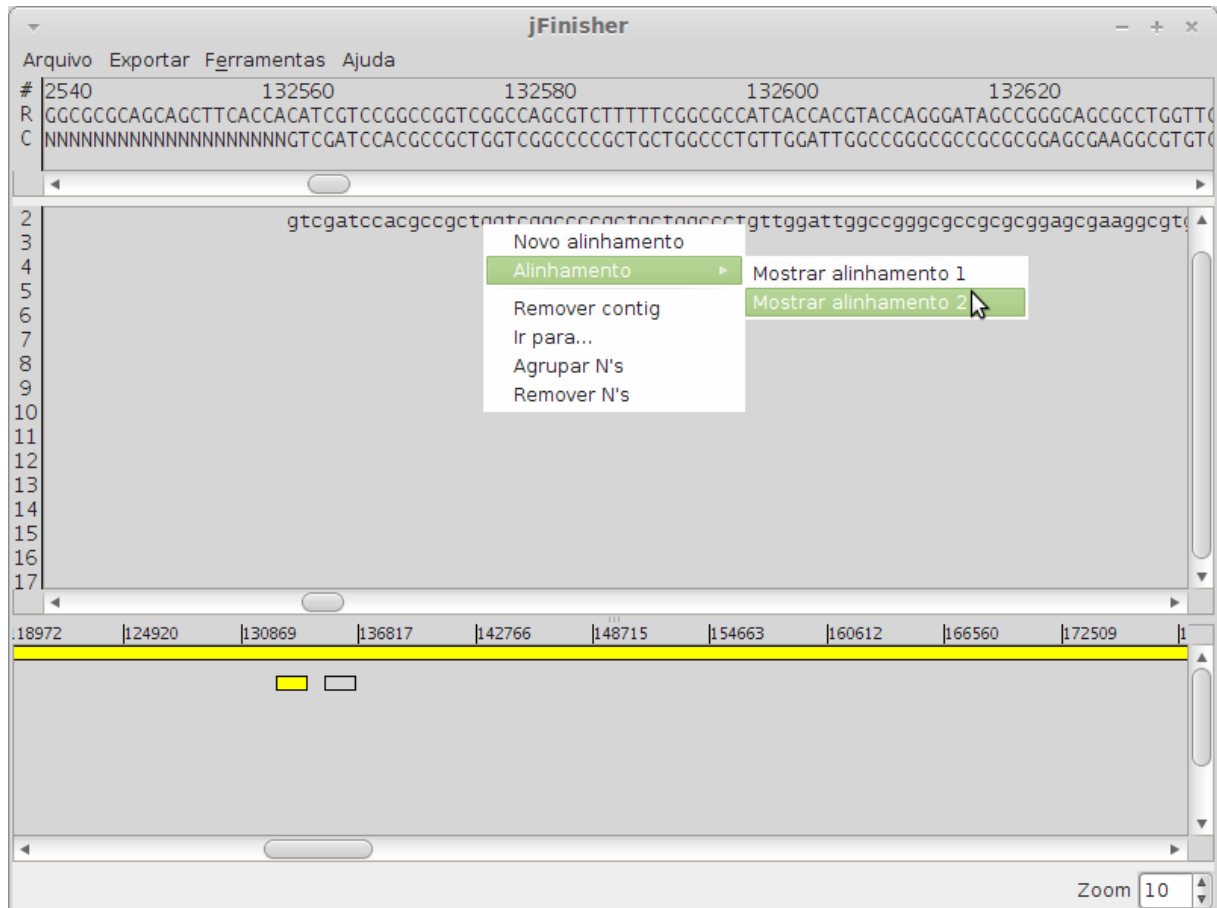


Figura 18 – Múltiplos alinhamentos por contig



É possível remover um alinhamento específico clicando com o botão direito sobre o mapa no alinhamento desejado e escolhendo a opção “Remover este alinhamento” (Figura 19).

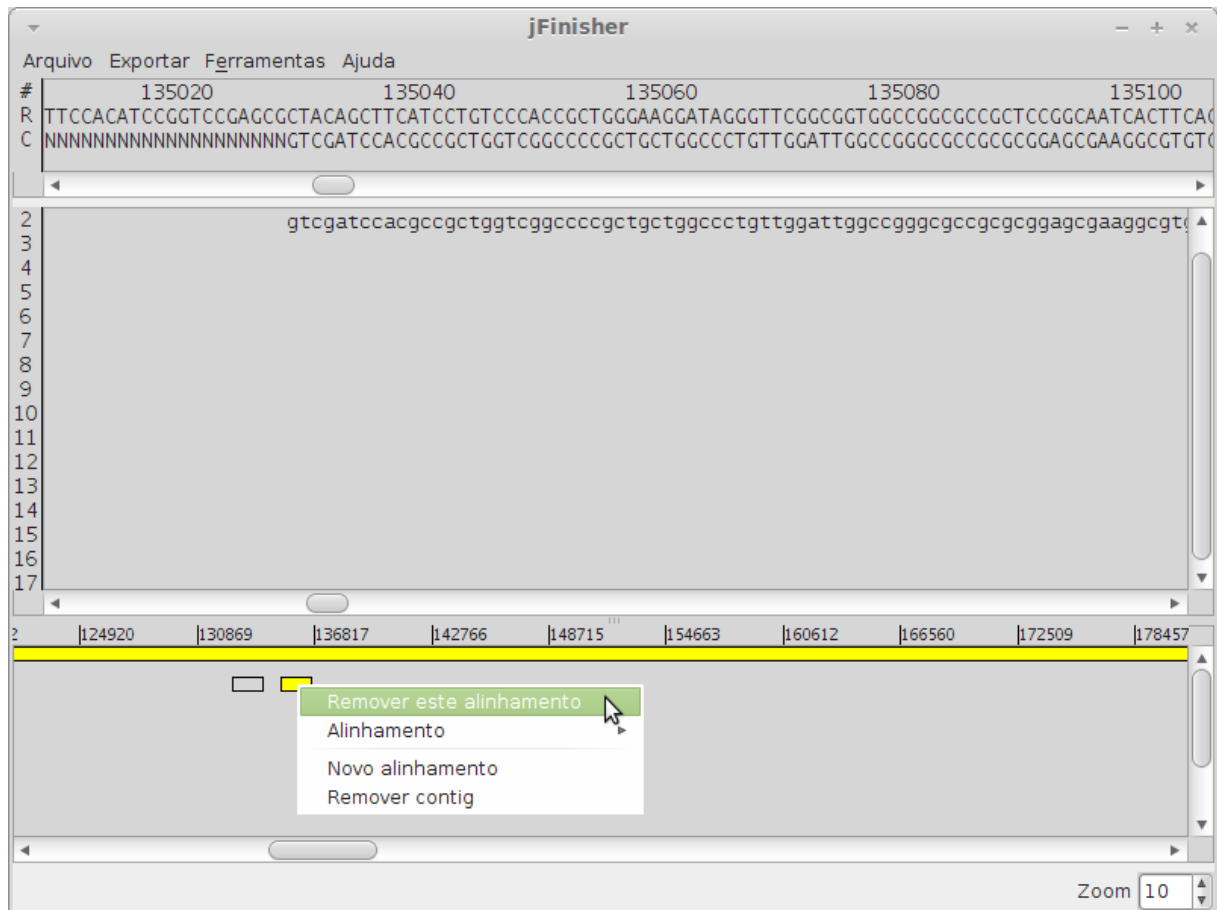


Figura 19 – Remover alinhamento

## 4.5 EDIÇÃO E LOCALIZAÇÃO

O jFinisher possui várias opções para edição das sequências:

- Remoção de um contig inteiro (e seus alinhamentos)
- Inserção ou remoção de determinados caracteres no contig
- Agrupamento ou remoção de caracteres “N”
- Copiar parte de sequências
- Localizar padrão ou navegar pela sequência manualmente

Todas estas opções são acessadas ao clicar com o botão direito sobre uma sequência (Figura 20).

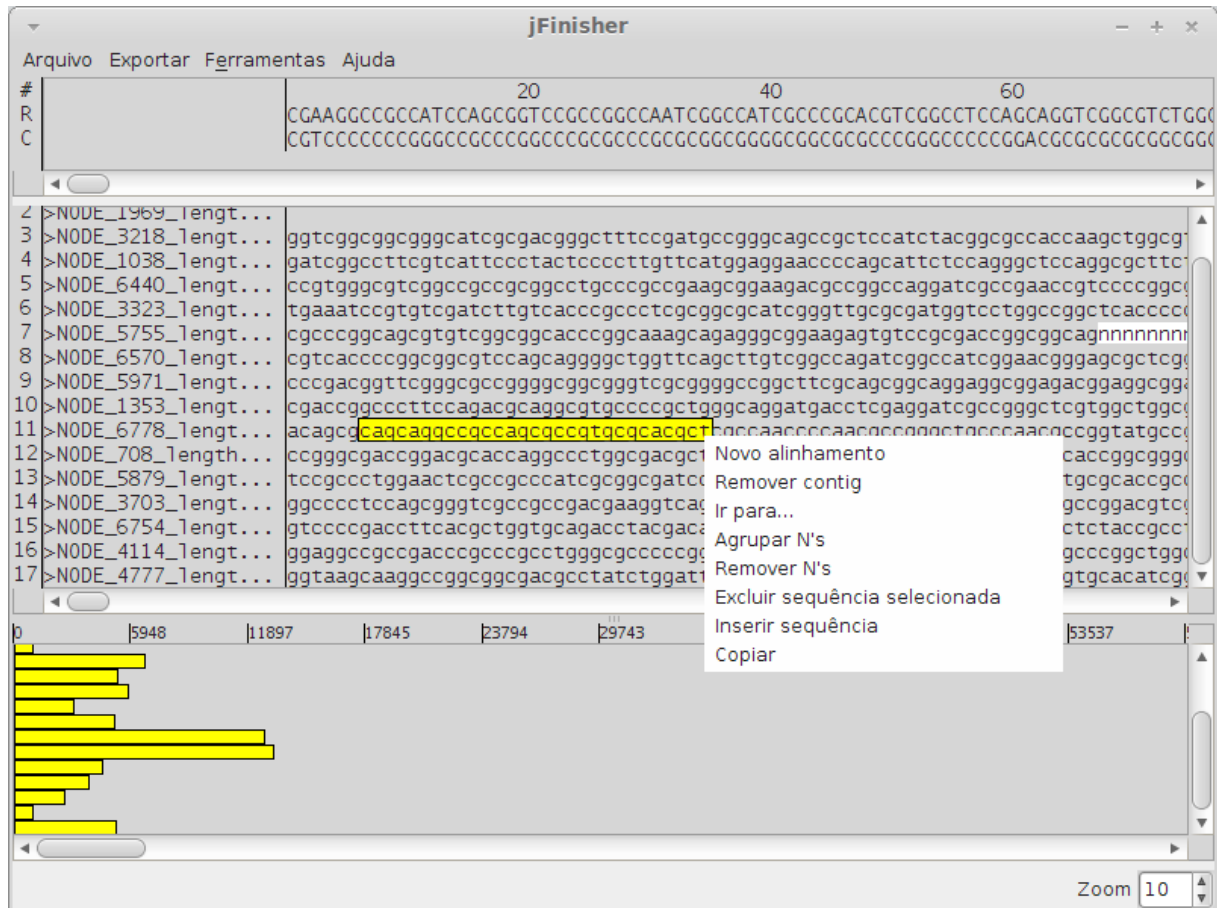


Figura 20 – Clique botão direito sobre sequência

A opção “Remover contig” quando acionada apresenta uma tela de confirmação antes de remover o contig selecionado do sistema (Figura 21).

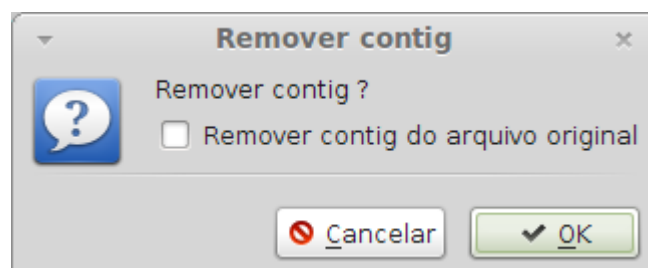


Figura 21 – Remover contig

Marcando a opção “Remover contig do arquivo original”, a sequência é removida do arquivo de origem (somente ocorre em arquivos multi-fasta).

“Excluir sequência selecionada” permite remover os caracteres selecionados do contig, após uma tela de confirmação. Esta ação é replicada no arquivo de origem.

“Inserir sequência” permite que dados sejam inseridos em certa posição do contig. Ao ser acionada esta opção ativa a tela de entrada de dados (Figura 22).

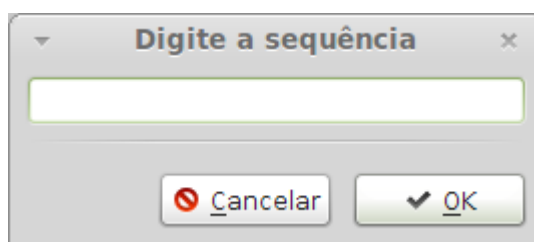


Figura 22 – Tela de entrada de dados

Ao confirmar a operação a sequência de caracteres informada é adiciona ao contig na posição selecionada. Esta operação altera o arquivo original.

As funções “Agrupar N’s” e “Remover N’s” tem papéis semelhantes. Ao selecionar uma das duas opções o sistema irá remover os caracteres “N” da sequência ou agrupar caracteres que estejam lado a lado em apenas um “N”.

“Copiar” transfere a seleção de caracteres da sequência para a área de transferência do sistema.

Para efetuar buscar e localização o sistema tem a opção de “Ir para” onde apresenta uma tela de navegação (Figura 23).

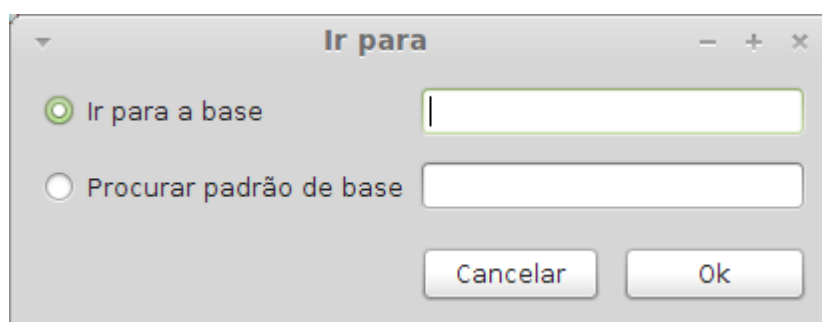


Figura 23 – Tela de navegação “Ir para”

“Ir para a base” dá a opção de navegar pela posição dos caracteres da sequência.

“Procurar padrão de base” busca uma sequência informada. Estas opções reposicionam a tela para a ocorrência, caso encontrada (Figura 24).

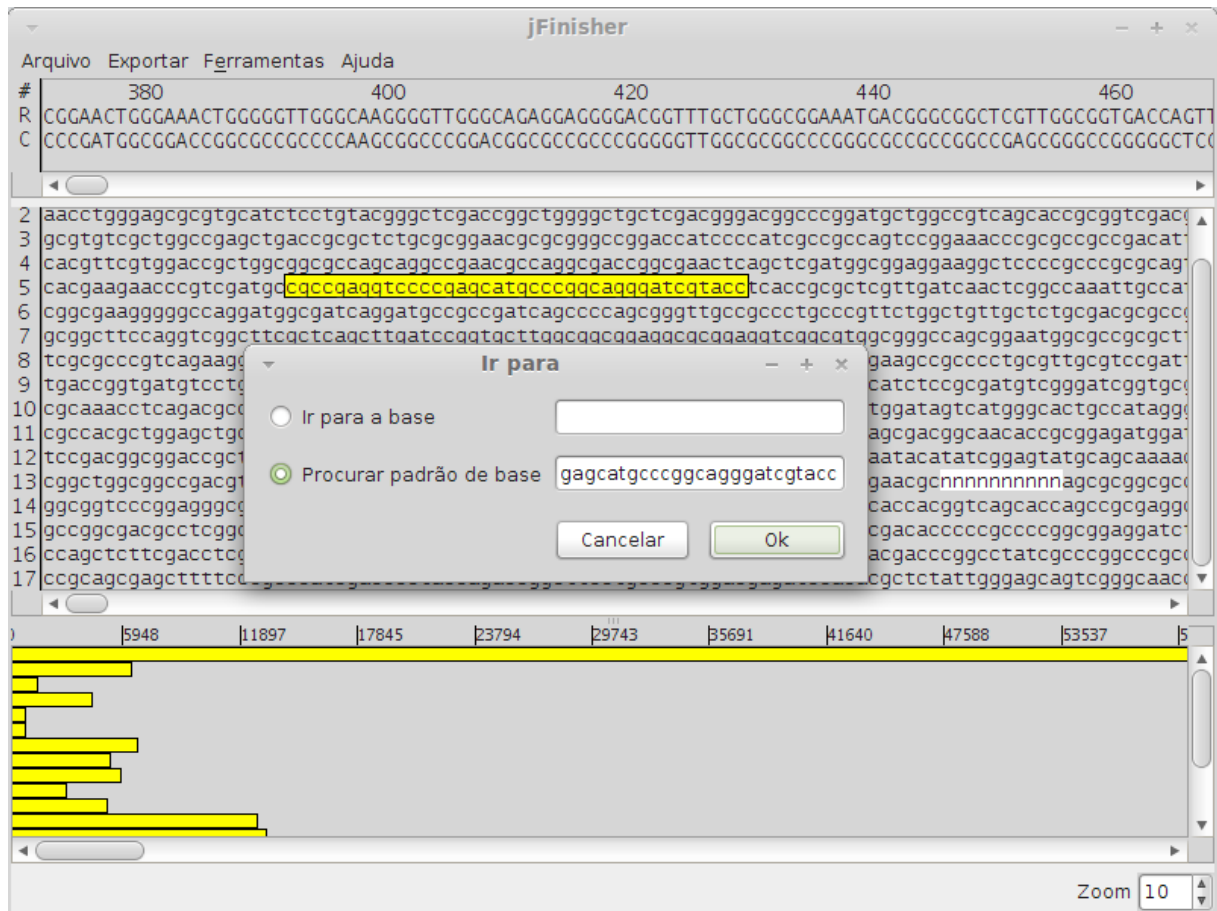


Figura 24 – Reposicionamento de tela pela ocorrência encontrada

## 4.6 PROJETO

O jFinisher permite que as alterações realizadas dentro do sistema sejam salvas. A opção “Salvar Projeto” no menu “Arquivo” permite a geração de um arquivo de projeto no formato XML que pode ser aberto posteriormente ou em outro computador.

A opção “Abrir Projeto” do menu “Arquivo” resgata um projeto previamente salvo com todas as alterações feitas. O arquivo de projeto não armazena as sequências e sim faz referência ao seu local de origem, sendo necessário sempre possuir os arquivos de sequências ao manipular projetos.

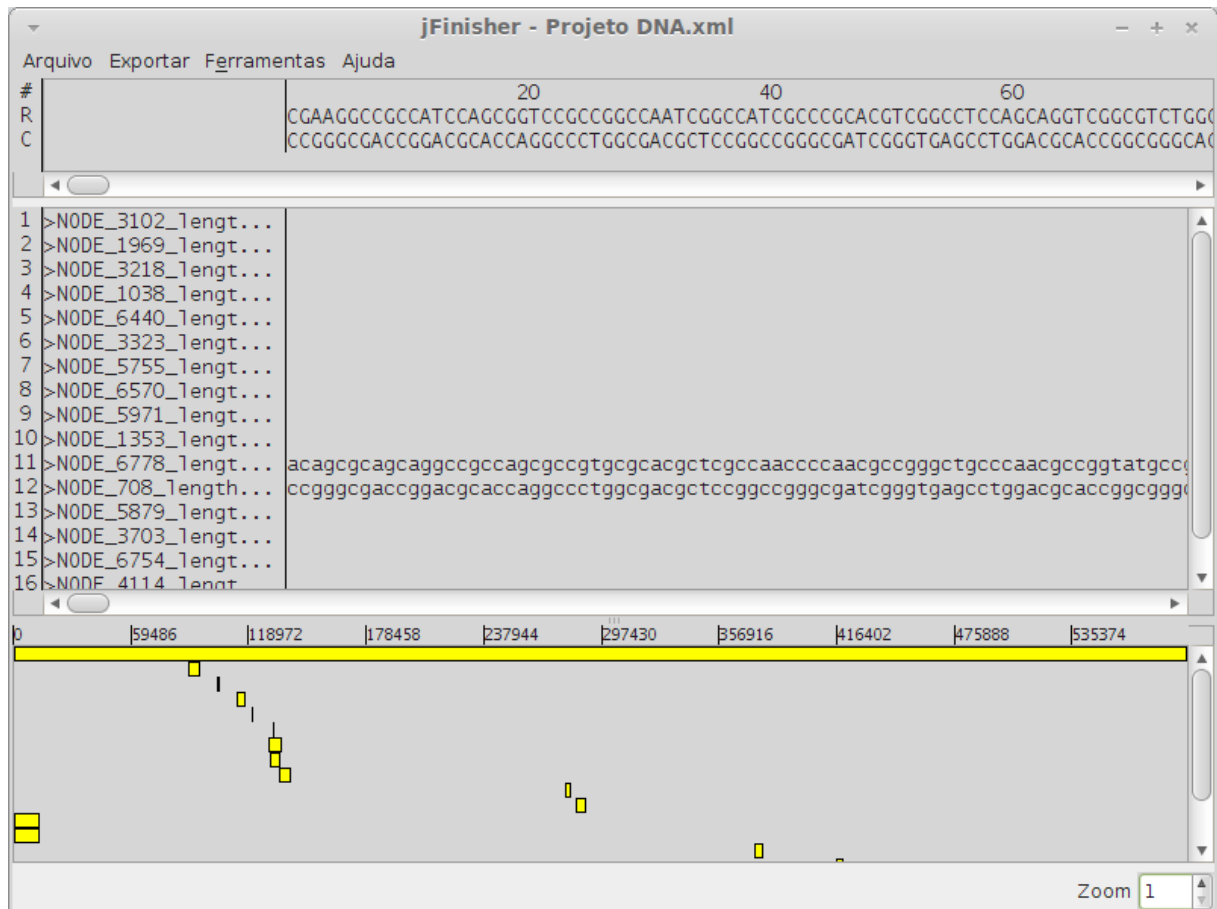


Figura 25 – Tela principal com projeto ativo

## 4.7 UNIR CONTIGS

A união de contigs pode ser realizada pelo menu “Ferramentas” opção “Unir seqüências”. Esta opção permite unificar arquivos distintos em uma só seqüência

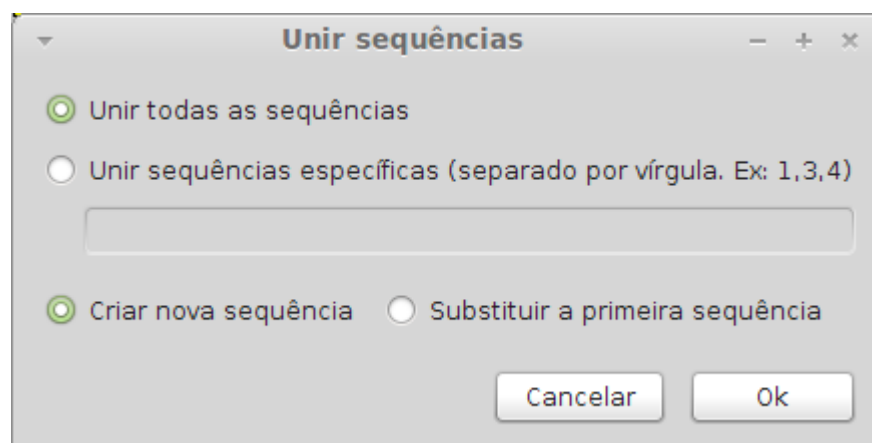


Figura 26 – Tela para unir seqüências

O usuário pode especificar quais contigs deseja unir, ou fazer uma união de todos. O arquivo pode sobrepor o primeiro contig (do conjunto de contigs que serão unidos) ou pode ser gerado um novo arquivo com a união.

Ao unir áreas sobrepostas entre os contigs o sistema realiza o consenso entre as áreas, utilizando como método de desempate o caractere que estiver no maior contig. Caso áreas na união não possuam sequências (Ex.: espaço entre dois contigs) o sistema irá preencher o espaço com o caractere “N”.

## 4.8 EXPORTAR

O resultado dos alinhamentos, edições e análises feitas pelo jFinisher podem ser exportadas. No menu “Exportar” (Figura 27) é possível salvar: o mapa de sequências, as sequências de contig ou a sequência consenso.

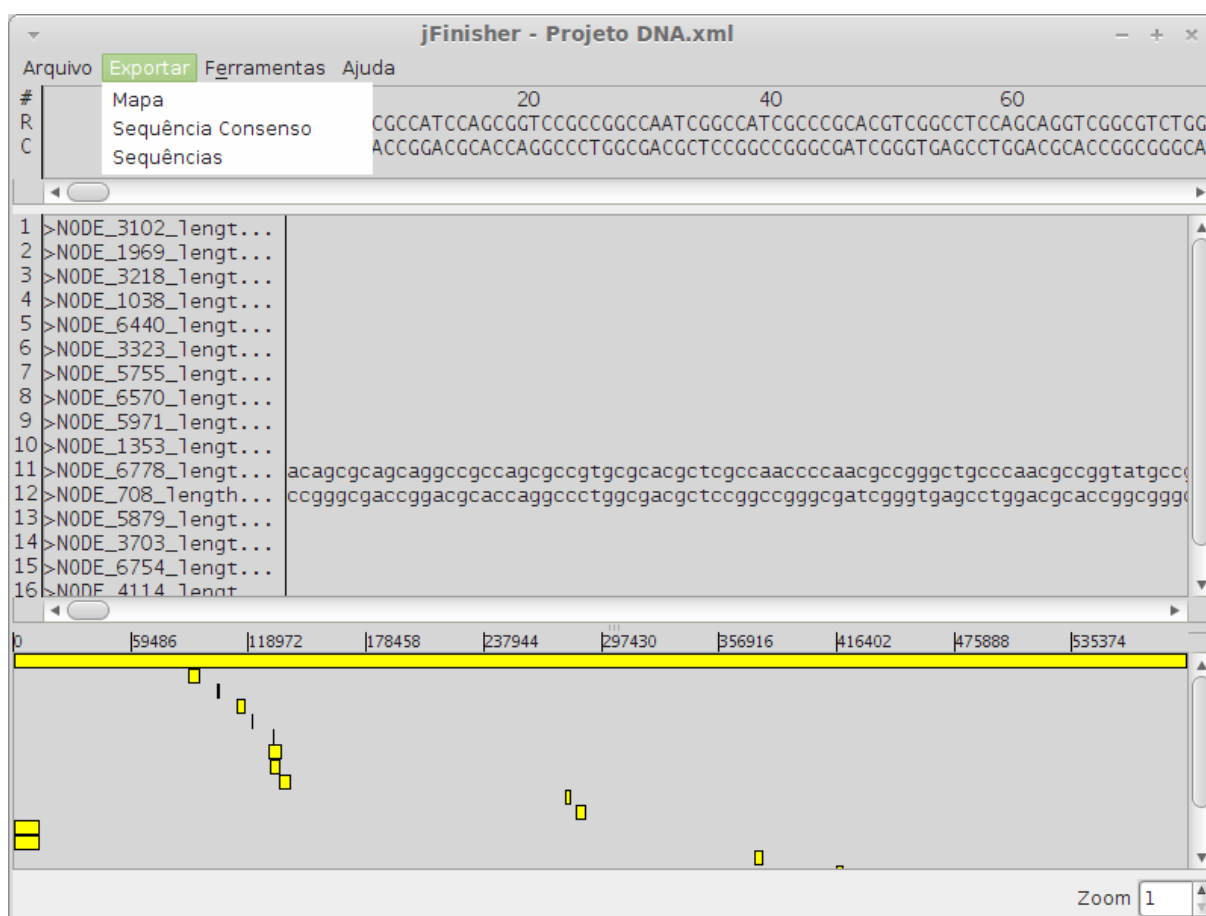


Figura 27 – Menu exportar

Em todas as opções de exportação, será mostrada uma tela para escolha do local e nome do arquivo que será gerado.

## 5 CONSIDERAÇÕES FINAIS

O planejamento e desenvolvimento do software JFinisher nos proporcionou um contato inicial e prático com a bioinformática. Ao realizar o projeto e aplicar os conceitos de alinhamento e edição de sequência foi possível entender uma parte dos problemas desta área.

Conseguimos criar uma ferramenta que unifica funcionalidades que são constantemente utilizadas por bioinformatas. Criamos uma nova interface visual, baseada nas necessidades para manipulação dos dados. O sistema teve foco na usabilidade e otimização e suporta arquivos grandes, de milhões de pares de base, algo pouco comum nos softwares com interface gráfica.

As funcionalidades de alinhamento conseguiram ser implementadas corretamente, porém com algumas limitações. Sequências muito grandes (acima de 15kb) não conseguem ser alinhadas, pois o método escolhido demanda muita memória e muitas vezes excede o limite do computador.

A edição das sequências de contigs foi facilitada pelo uso de bibliotecas que manipulam arquivos da área. Esta etapa pode ser expandida por novos recursos referentes à edição da sequência.

A opção de salvar e abrir projeto torna o programa mais flexível, podendo salvar seu próprio andamento, possibilitando que o trabalho possa ser feito em outros momentos ou repassado para outras pessoas. Funcionalidade que agrega grande valor ao produto.

As várias opções de exportar o trabalho efetuado fazem com que o resultado da aplicação do software possa ser utilizado em projetos reais, tendo a ferramenta como intermediária de um longo processo de análise e manipulação de sequências. Isto pode ser ampliado, possuindo mais opções para exportar arquivos.

Com este projeto, desenvolvemos uma ferramenta nova, que tem a capacidade de manipular e alinhar grandes sequências localmente, possibilitando ao usuário uma interface de fácil navegação e usabilidade. Esta ferramenta pode ser utilizada amplamente, em várias etapas na edição de sequências.

Durante o processo, várias ideias foram criadas, mas nem todas implementadas. Isso pelo fato do tempo de projeto, escopo e complexidade. Porém todas as ideias sugeridas foram analisadas. Com isto é possível uma ampliação do

software, tornando mais abrangente e com mais funções agregadas. Algumas funcionalidades e melhorias que podem ser implementadas são listadas abaixo:

- Alinhamento: utilizar outras técnicas para alinhamento em conjunto com as atuais, possibilitando que sequências grandes também possam ser alinhadas. Elas também podem melhorar o desempenho dos alinhamentos. Uma técnica que pode ser implementada é o algoritmo BLAST.
- Arquivos GBK como contigs: possibilitar que arquivos GenBank sejam abertos como contigs e que tenham as mesmas funcionalidades de edição e manipulação atuais dos arquivos Fasta.
- Sequências em linha: possibilitar que após múltiplos alinhamentos de uma mesma sequência ela seja replicada em linha. Este processo vai possibilitar os casos em que o contig se repete durante um genoma possam ser corretamente mapeados.
- Abrir mais formatos de arquivos: possibilitar a abertura e manipulação de sequências em outros formatos como: ColorSpace, FastaQ e outros mais utilizados na área.
- Mover sequência graficamente: dar a opção de alinhar manualmente a sequência através de entrada de valores ou graficamente com o mouse.
- Importar/Exportar alinhamentos BAM/SAM: implementar o controle destes arquivos que são um padrão para alinhamentos entre sequências.
- Comandos via console: habilitar sistema para alinhar sequências via console, sem utilização da interface gráfica
- Gerar gráficos que sejam úteis e auxiliem nas funções do programa
- Permitir que alinhamento seja feito baseado na sequência consenso
- Melhor interação entre sequências, efetuar ações múltiplas, selecionando número de sequências desejado
- Refinar sistema de erros de log
- Melhorar interface ao identificar arquivos por cores, com opções de personalização
- Estudar melhor divisão de sistemas por thread, segmentando operações que tem alto custo computacional



O software foi desenvolvido como um projeto acadêmico e, por decisão da equipe, liberado pela licença aberta GNU General Public License (GPL). O software e seu código fonte está disponível no repositório SourceForge e pode ser acessado através do link <http://sourceforge.net/projects/jfinisher/>

## REFERÊNCIAS

ALTSCHUL, S. F; et al. Basic local alignment search tool. **Journal of Molecular Biology**. v. 215, 1990, p. 403-410.

**ARTEMIS:** Genome Browser. Disponível em <<http://www.sanger.ac.uk/resources/software/artemis/>> Acesso em 10 de setembro de 2011.

MARCOS DA SILVA, Douglas; **UML:** Guia de consulta rápida. São Paulo: Novatec Editora, 2001.

**JAVA API.** Java™ Platform, Standard Edition 6 API Specification. Disponível em <<http://docs.oracle.com/javase/6/docs/api/>>. Acesso em 22 de janeiro de 2012.

JANDL JUNIOR, P. **JAVA:** Guia do Programador. São Paulo: Novatec Editora, 2007.

GUIZELINI, Dieval; **JContigSort.** Disponível em <<http://sourceforge.net/projects/jcontigsort/>>. Acesso em 22 de fevereiro de 2012.

MOUNT, David W. **Bioinformatics:** sequence and genome analysis. Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2004.

MOUSTAFA, Ahmed. **JAligner:** Open source Java implementation of Smith-Waterman. Disponível em <<http://jaligner.sourceforge.net>> Acesso em 22 de fevereiro de 2012.

NEEDLEMAN, S. B.; WUNSCH, C. D. A. General Method Applicable to the Search for Similarities in the Amino Acid Sequence of Two Proteins. **Journal of Molecular Biology**, 1970, v. 48, p.443-453.

PRESSMAN, Roger S. **Engenharia de Software** - 6. ed. São Paulo: Makron Books, 2006.

PENA, Sergio D. J. Era Pós-Genômica - anno V, 2005. **Scientific American Brasil** Disponível em <[http://www2.uol.com.br/sciam/artigos/era\\_pos-genomica\\_-\\_anno\\_v.html](http://www2.uol.com.br/sciam/artigos/era_pos-genomica_-_anno_v.html)> Acesso em 10 de março de 2012.

PROSDOCIMI, F. et al. Bioinformática: Manual do Usuário. **Biotecnologia Ciência & Desenvolvimento**. n. 29, 2002.

PROSDOCIMI, F. **Introdução À Bioinformática**. Disponível em <[http://www2.bioqmed.ufrj.br/prosdocimi/chicopros/FProsdocimi07\\_CursoBioinfo.pdf](http://www2.bioqmed.ufrj.br/prosdocimi/chicopros/FProsdocimi07_CursoBioinfo.pdf)>. Acesso em 05 de janeiro de 2012.

SMITH, T. F.; WATERMAN, M. S.; Identification of Common Molecular Subsequences. **Journal of Molecular Biology**, 1981, v. 147, p. 195-197.

WALNES, Joe; SCHAIBLE, Jörg; et al. **XStream**. Disponível em <<http://xstream.codehaus.org/>>. Acesso em 22 de fevereiro de 2012.

WATSON, James D.; BAKER, Tania A.; BELL, Stephen P.; GANN, Alexander; LEVINE, Michael, LOSICK, Richard. **Biologia Molecular do Gene** – 5. ed. Porto Alegre: Atrmed, 2006.

**WBS TOOL**. Software web gratuito para criação e edição de WBSs. Disponível em <<http://www.wbstool.com/>>. Acesso em 29 de janeiro de 2012.

ZERBINO, D. R.; BIRNEY, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. **Genome Research**, v.18, p. 821-829, 2008. doi:10.1101/gr.074492.107.

## APÊNDICE

### 1 DIAGRAMA DE CASO DE USO

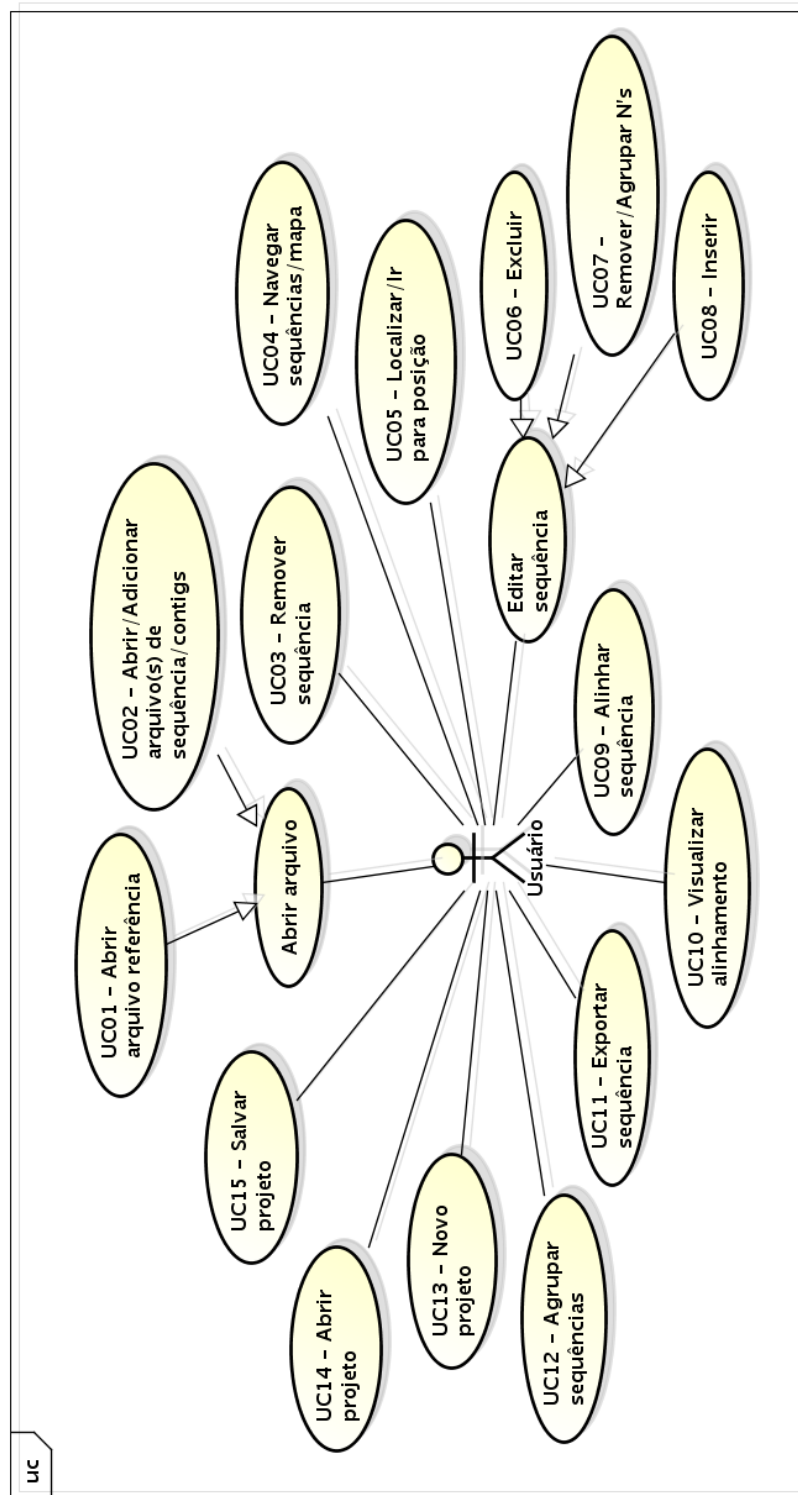


Figura 28 – Diagrama de Caso de uso

## 2 DIAGRAMA DE CLASSES

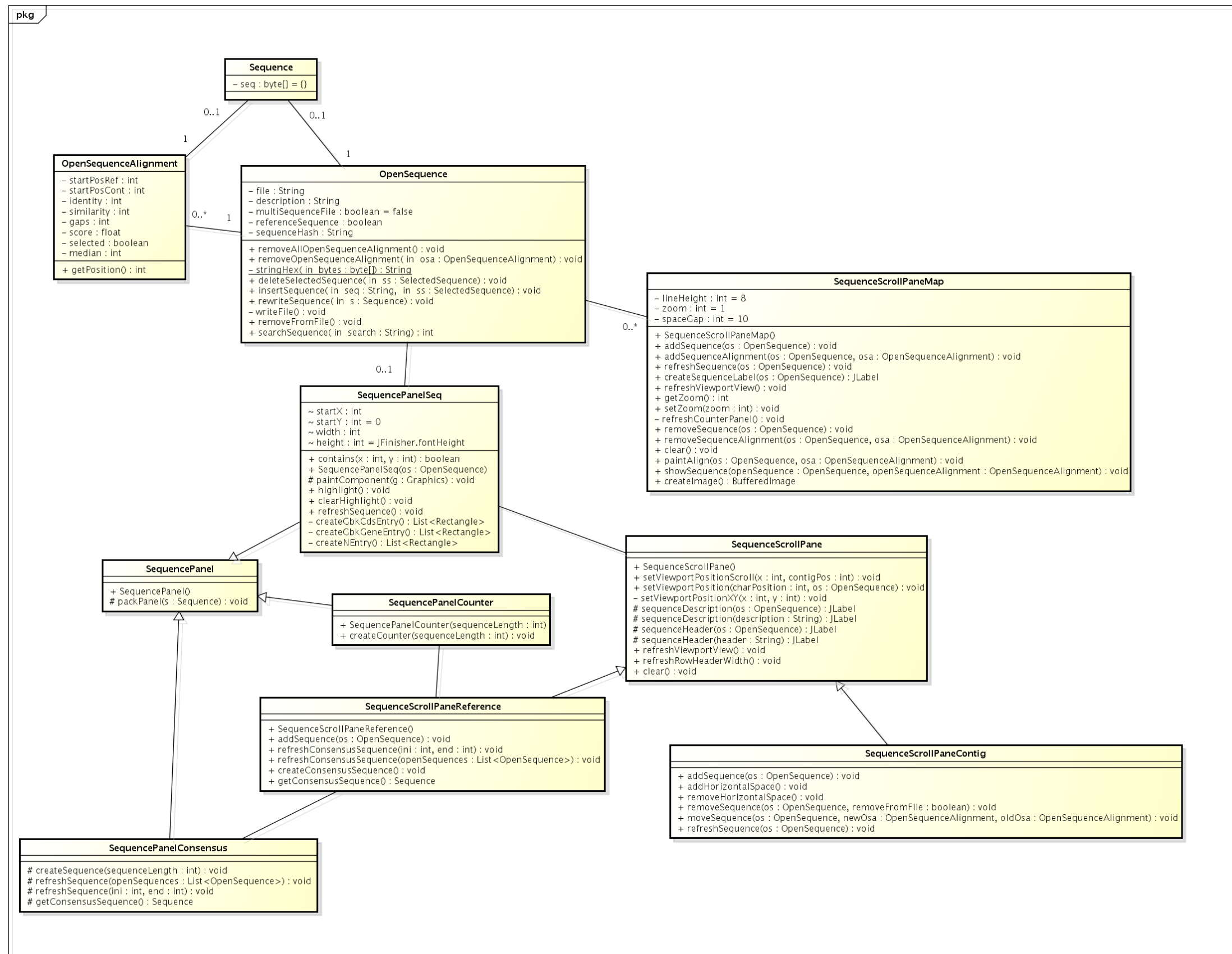


Figura 29 – Diagrama de Classe

### 3 ESPECIFICAÇÃO DE CASOS DE USO

UC01 – Abrir arquivo referência

#### Descrição

Este caso de uso descreve a abertura de um arquivo de sequência que será utilizado como referência para as operações do software.

#### Pré-condições

1. O sistema deve estar iniciado
2. Não possuir nenhum arquivo ou projeto aberto

#### Pós-condições

1. Abrir e mostrar na tela a sequência de referência, contador e sequência consenso
2. Gerar o mapa de sequências com o contador próprio

#### Ator Primário

Usuário

#### Fluxo de Eventos Principal

1. O sistema apresenta a tela (**DV1**)
2. O usuário seleciona a opção “Arquivo, Abrir Referência” (**DV2**)
3. O sistema apresenta a tela de seleção de arquivo (**DV3**)
4. O usuário seleciona o arquivo de referência (**R1**) (**A1**) (**E1**)
5. O sistema abre o arquivo de referência e mostra na tela (**R2**) (**R3**) (**R4**) (**R5**) (**R6**) (**R7**)
6. O caso de uso é finalizado

#### Fluxos Alternativos

**A1:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela de seleção de arquivos

2. O caso de uso é reiniciado

### **Fluxos de Exceção**

**E1.** Arquivo selecionado não é reconhecido:

1. O apresenta a mensagem de erro (**DV4**)
2. O caso de uso é reiniciado

### **Regras de Negócio**

**R1:** O sistema somente aceita arquivos de sequência no formato fasta ou genbank, com as extensões "gbk", "fas", "fa", "fasta" ou "fna", podendo ser selecionado apenas um arquivo.

**R2:** O sistema mostra na parte superior do programa a sequência aberta, com opção de rolagem lateral.

**R3:** O sistema gera um contador acima da sequência aberta que irá identificar a posição dos caracteres.

**R4:** Gerar um mapa gráfico na parte inferior do programa onde a sequência referência seja o total máximo da tela (100%), utilizando o tamanho da sequência aberta como parâmetro de proporção entre sequências abertas posteriormente.

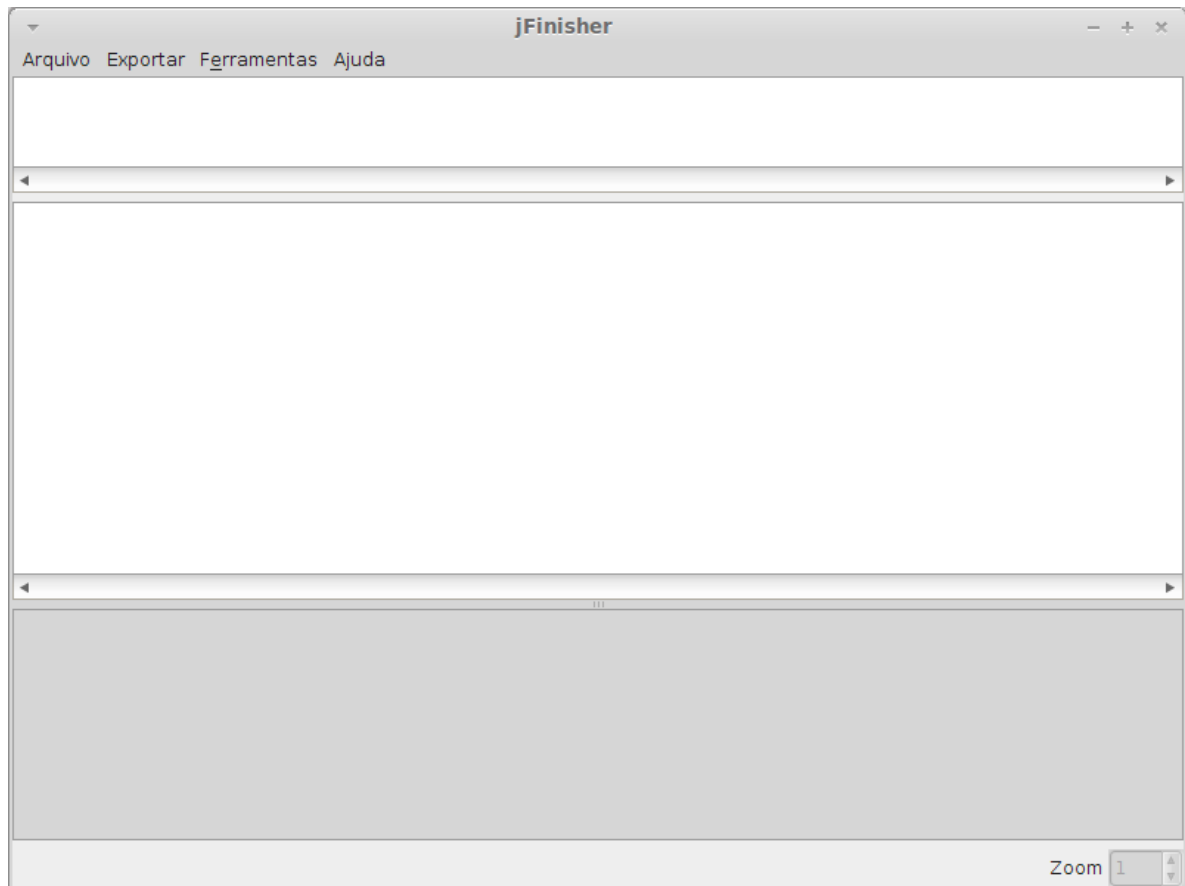
**R5:** O sistema gera uma segunda sequência abaixo da sequência referência (sequência consenso), que será atualizada conforme outras sequências.

**R6:** Opções do menu principal que estão bloqueadas são liberadas para acesso. Campo "Zoom" é liberado para alteração. Opção "Arquivo, Abrir Referência" é bloqueada.

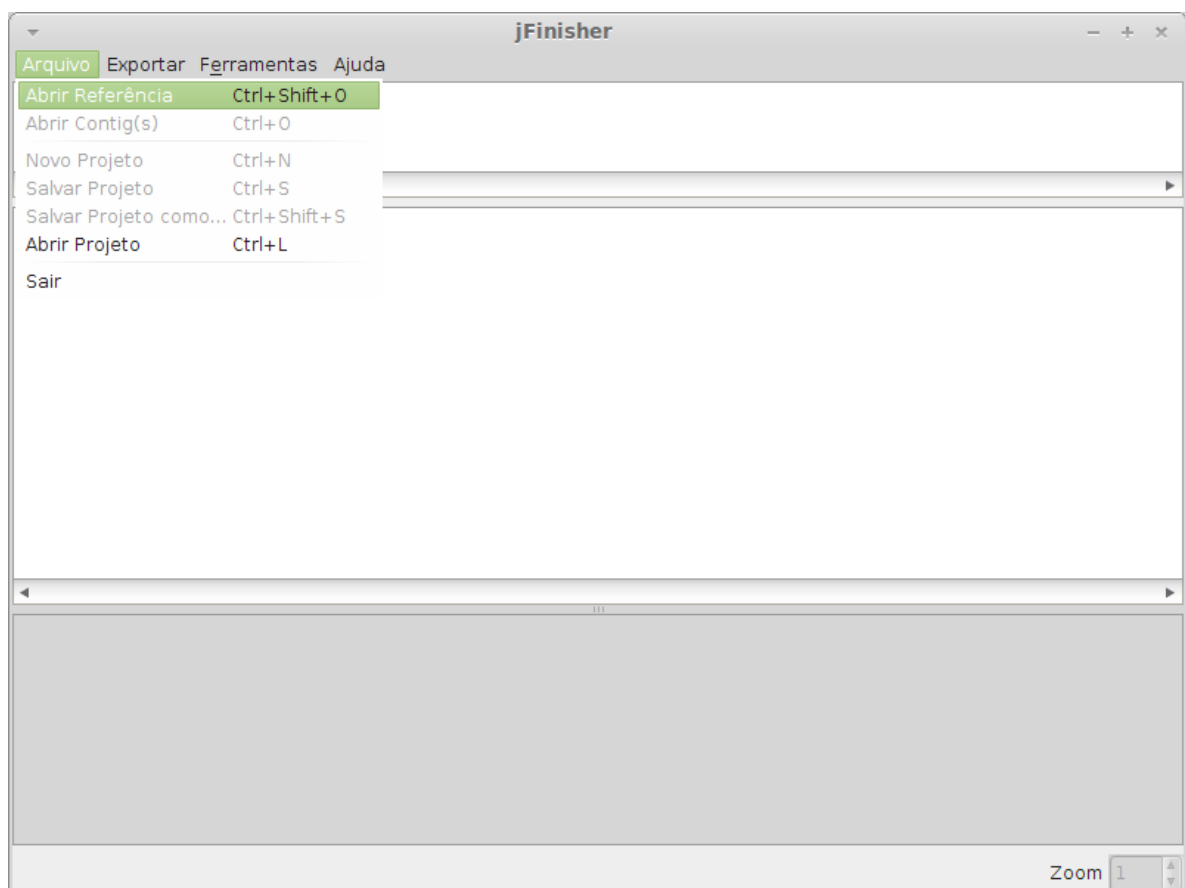
**R7:** Zoom do mapa tem como valor mínimo "1", que representa o maior tamanho possível na tela sem barra de rolagem. Ao aumentar este valor, o sistema deve aumentar o tamanho do mapa e todos seus componentes internos, mantendo a proporção entre eles.

### **Data View**

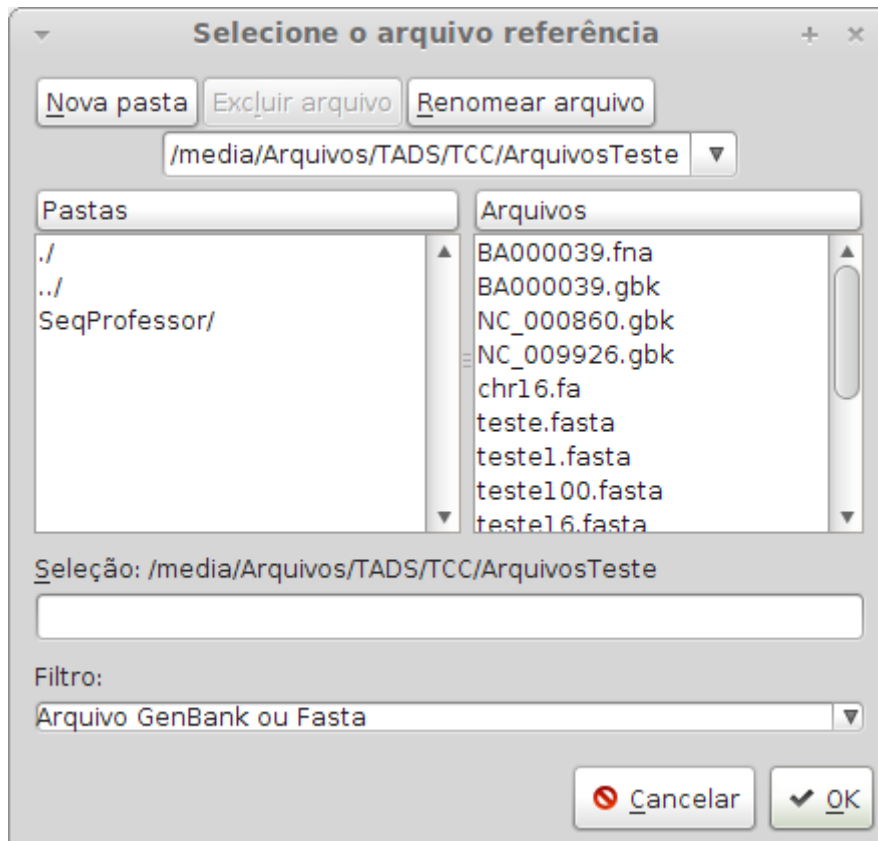
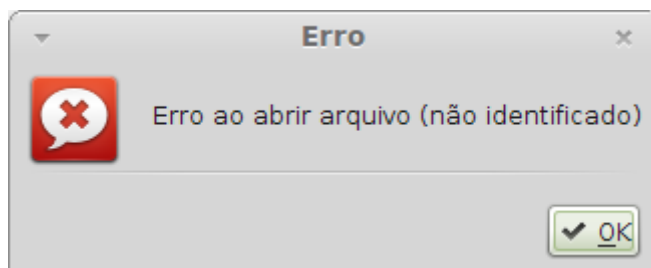
**DV1** – Tela principal do programa



## DV2 – Abrir referência





**DV3 - Selecionar arquivo****DV4 – Erro abrir arquivo****UC02 – Abrir-Adicionar arquivo(s) de sequência-contigs****Descrição**

Este caso de uso descreve a abertura de um ou mais arquivos de sequência (também chamados de contig) que serão analisados e manipulados pelo programa.

**Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto

### **Pós-condições**

1. Ter aberto e mostrado na tela as sequências/contigs
2. Adicionado no mapa as sequências

### **Ator Primário**

Usuário

### **Fluxo de Eventos Principal**

1. O sistema apresenta a tela com referência carregada (**DV1**)
2. O usuário seleciona a opção “Arquivo, Abrir Contig(s)” (**DV2**)
3. O sistema apresenta a tela de seleção de arquivo (**DV3**)
4. O usuário seleciona um ou mais arquivos (**R1**) (**A1**) (**E1**)
5. O sistema abre os arquivos selecionados e mostra na tela (**R2**) (**R3**)
6. O usuário clica sobre uma sequência (**R4**) (**A2**) (**A4**) (**A8**) (**A9**)
7. O usuário clica com o botão direito sobre a sequência (**R5**)
8. O usuário seleciona a opção “Ir para...” do menu gerado (**R6**) (**A3**) (**A5**) (**A6**) (**A7**)
9. O caso de uso é finalizado

### **Fluxos Alternativos**

**A1:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela de seleção de arquivos
2. O caso de uso é reiniciado

**A2:** Usuário clica sobre a sequência e arrasta cursor:

1. Sistema deve selecionar e marcar na tela a sequência que está sendo clicada e arrastada, caso atinja o limite da tela barra de rolagem deve acompanhar cursor do mouse
2. O caso de uso é finalizado

**A3:** Opção “Copiar” pressionada:

1. Sistema copia texto (**R7**)
2. O caso de uso é finalizado

**A4:** Alterado valor do campo “Zoom”:

1. Sistema recalcula tamanho do mapa (**R8**)

2. O caso de uso é finalizado

**A5:** Opção “Novo alinhamento” pressionada:

1. Sistema executa caso de uso “**UC09 - Alinhar sequência**”
2. O caso de uso é finalizado

**A6:** Opção “Remover contig” pressionada:

1. Sistema executa caso de uso “**UC03 - Remover sequência**”
2. O caso de uso é finalizado

**A7:** Opção “Agrupar N’s”, “Remover N’ s”, “Excluir sequência selecionada” ou “Inserir sequência” pressionada:

1. Sistema executa caso de uso “**Editar sequência (UC06, 07 e 08)**”
2. O caso de uso é finalizado

**A8:** Opção do menu “Ferramentas, Alinhar todos os contigs” pressionada:

1. Sistema executa caso de uso “**UC09 - Alinhar sequência**”
2. O caso de uso é finalizado

**A9:** Opção do menu “Ferramentas, Unir sequências” pressionada:

1. Sistema executa caso de uso “**UC12 - Agrupar sequências**”
2. O caso de uso é finalizado

## Fluxos de Exceção

**E1:** Arquivo selecionado não é reconhecido:

1. O apresenta a mensagem de erro (**DV4**)
2. O caso de uso é reiniciado

## Regras de Negócio

**R1:** O sistema somente aceita arquivos de sequência no formato fasta ou multi-fasta, com as extensões "fas", "fa", "fasta", "fna", "ffn", "faa", podendo ser selecionado um ou mais arquivos.

**R2:** O sistema mostra na parte central do programa as sequências abertas, com opção de rolagem lateral, uma sequência por linha, alinhadas a esquerda. Antes do início de cada sequência deve ser criado um contador sequencial que irá identificar a sequência, além de ter um espaço pré-determinado para descrição da sequência que será identificada por arquivo.

**R3:** Adiciona ao mapa na parte inferior do programa uma sequência por linha de maneira gráfica. A proporção de tamanho em relação à sequência de referência é dada pelo tamanho (quantidade de caracteres) de cada sequência.

**R4:** Sistema deve identificar sequência que foi clicada e marcar na tela o caractere selecionado.

**R5:** Sistema gera menu no local onde foi clicado com as opções:

“Novo alinhamento”

“Remover contig”

“Ir para...”

“Agrupar N’s”

“Remover N’ s”

Caso haja seleção de um ou mais caracteres o menu deve possuir também as opções:

“Excluir sequência selecionada”

“Inserir sequência”

“Copiar”

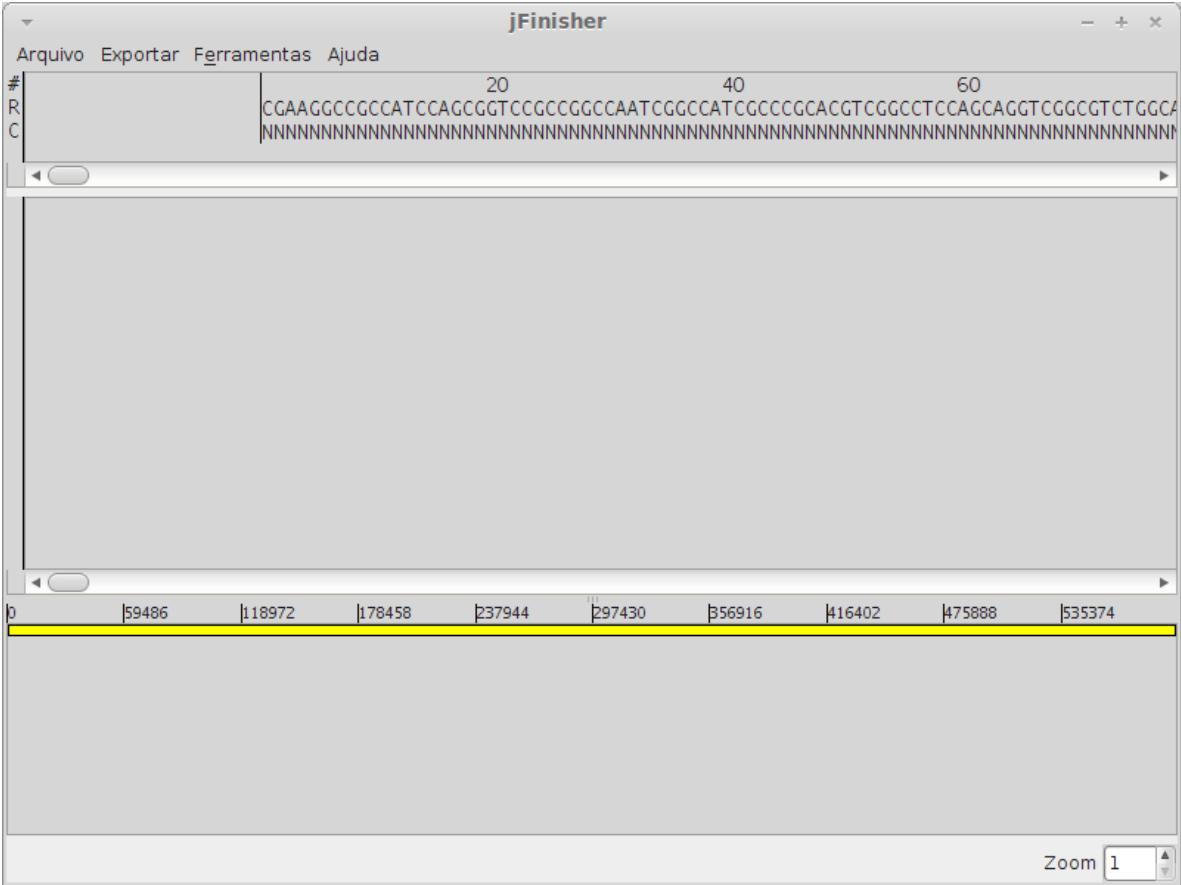
**R6:** Sistema executa caso de uso “**UC05 - Localizar/Ir para posição**”.

**R7:** Ao clicar em “Copiar”, o conteúdo selecionado na tela deve ir para a área de transferência do sistema operacional.

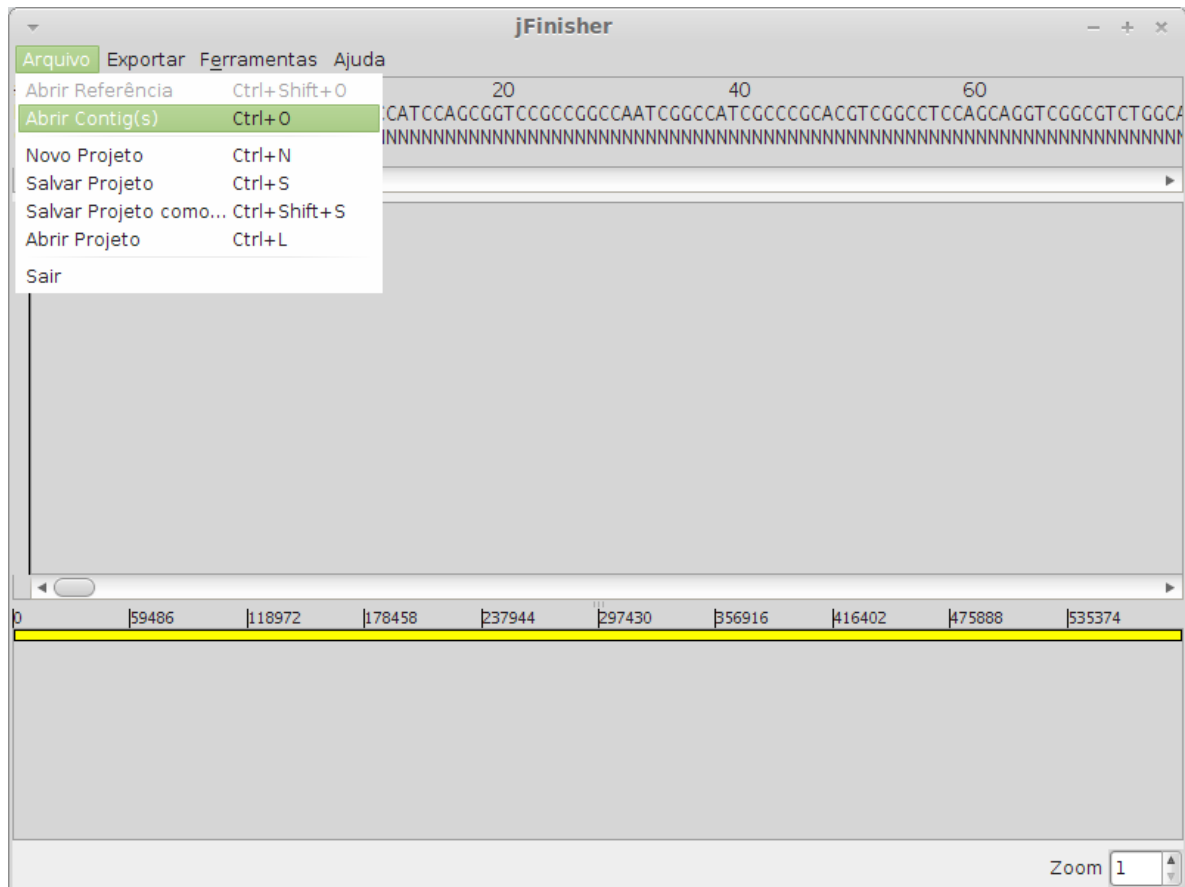
**R8:** Zoom do mapa tem como valor mínimo “1”, que representa o maior tamanho possível na tela sem barra de rolagem. Ao aumentar este valor, o sistema deve aumentar o tamanho do mapa e todos seus componentes internos, mantendo a proporção entre eles.

## **Data View**

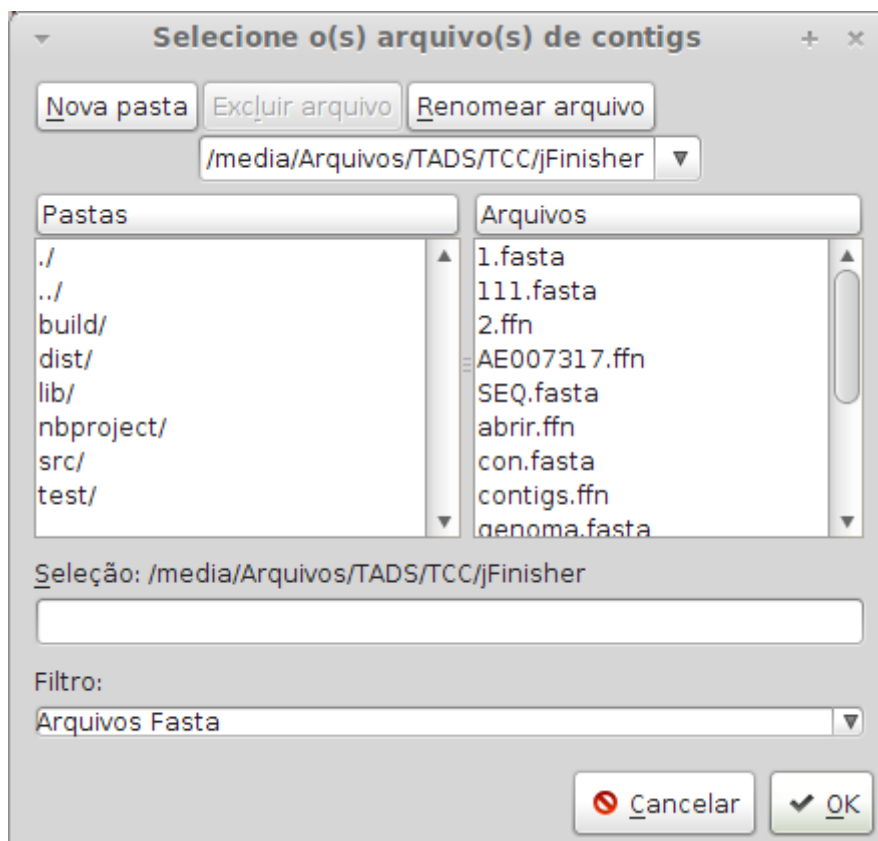
**DV1 –** Tela principal do programa com referência aberta

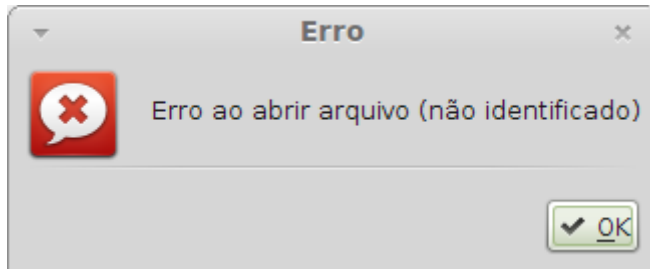


## DV2 – Abrir contig(s)



## DV3 – Selecionar arquivo(s)



**DV4 – Erro abrir arquivo****UC03 – Remover sequência****Descrição**

Este caso de uso descreve a remoção de uma sequência

**Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

**Pós-condições**

1. Ter removido a sequência desejada

**Ator Primário**

Usuário

**Fluxo de Eventos Principal**

1. O usuário clica com o botão direito em uma sequência e escolhe a opção “Remover contig” (**R1**)
2. O sistema apresenta a tela de confirmação (**DV1**)
3. O usuário clica no botão “Ok” (**A1**)
4. O sistema remove a sequência (**R2**)
5. O caso de uso é finalizado

**Fluxos Alternativos**

**A1:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela de confirmação sem efetuar nenhuma ação
2. O caso de uso é finalizado

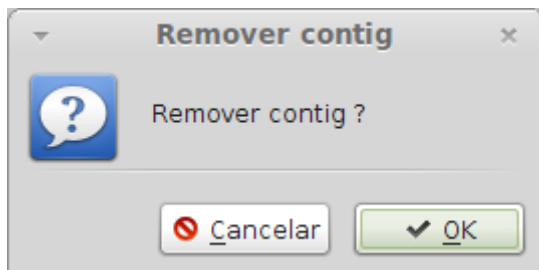
### **Regras de Negócio**

**R1:** Somente sequências contigs podem ser removidas (todas menos referência).

**R2:** O sistema remove da tela de sequências e do mapa gráfico as referências ao contig selecionado para remoção.

### **Data View**

**DV1 – Confirmação**



UC04 – Navegar sequências-mapa

### **Descrição**

Este caso de uso descreve a navegação pela interface do software

### **Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

### **Pós-condições**

1. Ter possibilitado a navegação e visualização das informações dos arquivos graficamente

### **Ator Primário**

Usuário



### **Fluxo de Eventos Principal**

1. O usuário arrasta a barra de rolagem horizontal na tela principal (**DV1**) (**R1**)
2. O usuário clica sobre uma sequência no mapa
3. O sistema reposiciona a tela para visualização da sequência clicada (**R2**)
4. O usuário clica duas vezes sobre uma sequência no mapa
5. O sistema marca a sequência na tela (**R3**)
6. O usuário desmarca a opção “Ferramentas, Travar barras” (**DV2**) (**R4**)
7. O usuário arrasta a barra de rolagem horizontal (**R5**)
8. O caso de uso é finalizado

### **Regras de Negócio**

**R1:** O sistema deve mover todas as 3 partes do sistema ao mesmo tempo (proporcionalmente) para que as sequências fiquem alinhadas conforme abertas.

**R2:** A tela central (sequências de contigs) deve ficar centralizada sem perder a referência de posição com outras sequências. Deve ser possível visualizar o início da sequência no quadro central referente à sequência clicada no mapa.

**R3:** O sistema executar a regra **R2** e também marca toda a sequência no quadro central.

**R4:** O sistema deve desvincular as barras relacionadas pela regra **R1**.

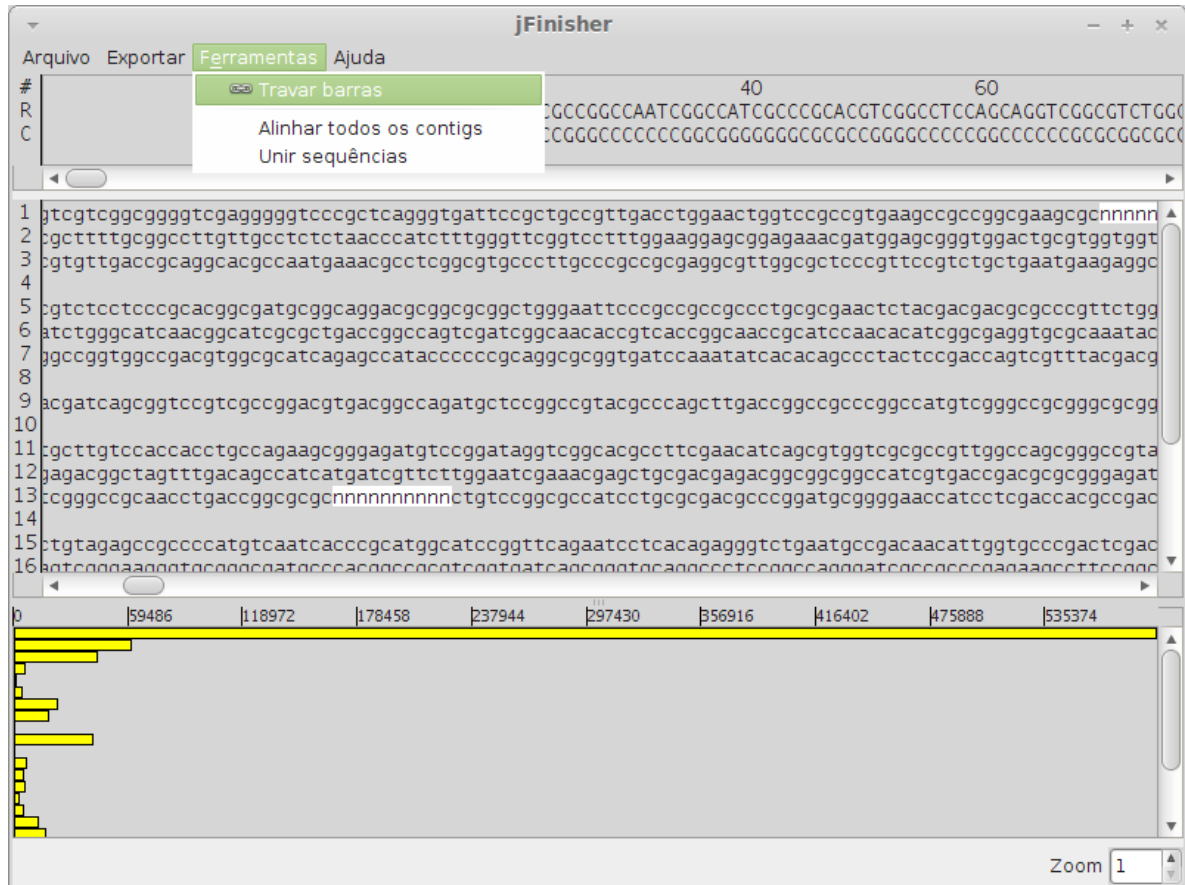
**R5:** Cada quadro deve mover-se independentemente do outro.

### **Data View**

**DV1** – Tela principal do programa



## DV2 – Travar barras



## UC05 – Localizar-lr para posição

### Descrição

Este caso de uso descreve a localização de caracteres dentro das sequências abertas no sistema.

### Pré-condições

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

### Pós-condições

1. Ter realizado a localização/posicionamento desejado

## **Ator Primário**

Usuário

### **Fluxo de Eventos Principal**

1. O usuário clica com o botão direito sobre uma sequência aberta (referência ou contig) e seleciona a opção “Ir para...” (**DV1**)
2. O sistema apresenta a tela de localização (**DV2**) (**R1**)
3. O usuário informa um número e clica em Ok (**R2**) (**A1**) (**E1**)
4. O sistema posiciona a tela conforme ocorrência
5. O usuário clica em Cancelar
6. O caso de uso é finalizado

### **Fluxos Alternativos**

**A1:** Usuário marca “Procurar padrão de base”:

1. O usuário digita uma sequência e clica em Ok (**R3**) (**E1**)
2. O sistema posiciona a tela conforme ocorrência
3. O usuário clica em Cancelar
4. O caso de uso é finalizado

### **Fluxos de Exceção**

**E1:** Ocorrência não encontrada:

1. O sistema não executa nenhuma ação
2. O caso de uso é finalizado

### **Regras de Negócio**

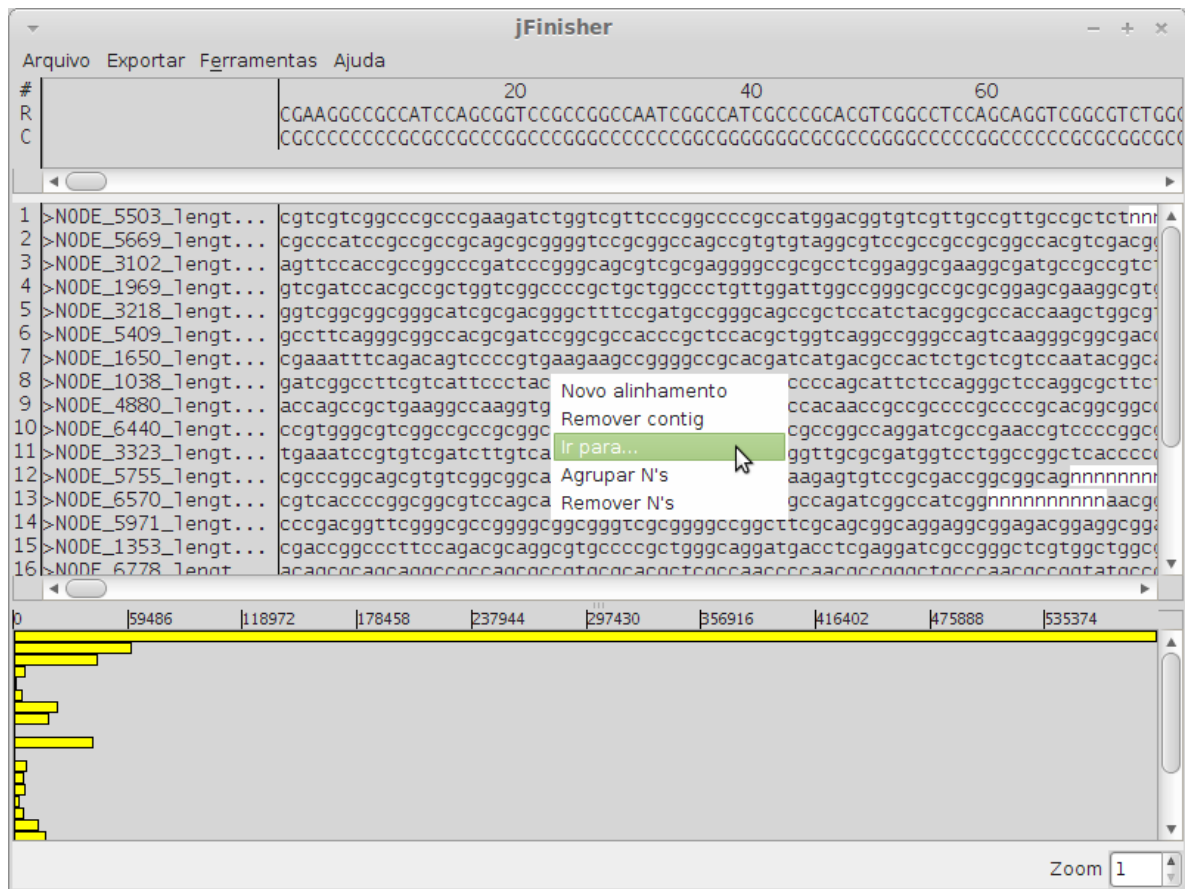
**R1:** O sistema marca automaticamente a opção “Ir para a base” e o cursor fica posicionado no campo de entrada ao lado.

**R2:** O sistema deve localizar a posição informada na sequência onde está sendo efetuada a localização e posicionar na tela a ocorrência, marcando o caractere encontrado.

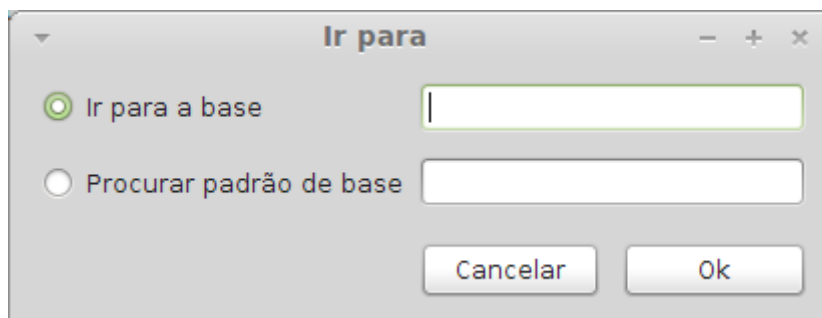
**R3:** O sistema deve localizar a sequência informada na sequência onde está sendo efetuada a localização e posicionar na tela a ocorrência, marcando o caractere encontrado.

## Data View

### DV1 – Seleção “Ir para” no menu



### DV2 – Tela de localização



### UC06 – Editar sequência – Excluir

## Descrição

Este caso de uso descreve a exclusão de parte de sequências

**Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

**Pós-condições**

1. Ter excluído a sequências selecionada, replicando a alteração no arquivo

**Ator Primário**

Usuário

**Fluxo de Eventos Principal**

1. O usuário clica e marca uma sequência (**R1**)
2. O usuário clica com o botão direito sobre a seleção e escolhe a opção “Excluir sequência selecionada” (**R2**)
3. O sistema apresenta a tela de confirmação de exclusão (**DV1**)
4. O usuário clica em Ok (**A1**)
5. O sistema remove a sequência selecionada (**R3**)
6. O caso de uso é finalizado

**Fluxos Alternativos**

**A1:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela de confirmação sem realizar nenhuma ação.
2. O caso de uso é finalizado.

**Regras de Negócio**

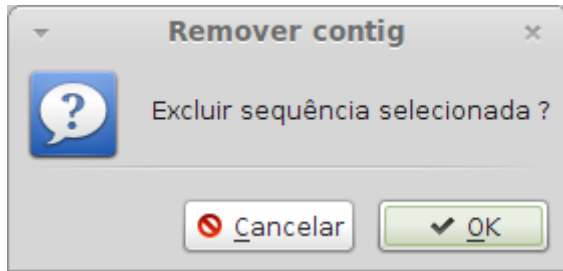
**R1:** O sistema deve marcar na tela com destaque a sequência selecionada pelo mouse.

**R2:** O sistema gera um menu no local onde foi realizado o clique com esta opção.

**R3:** O sistema remove a sequência selecionada e atualiza seus dados no programa e no arquivo em que a sequência se encontra armazenada.

**Data View**

**DV1** – Tela de confirmação



## UC07 – Editar sequência - Remover-Agrupar Ns

### Descrição

Este caso de uso descreve a exclusão automática ou agrupamento de sequências com caractere “N”

### Pré-condições

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

### Pós-condições

1. Ter agrupado/excluído os caracteres “N” de uma sequência, replicando a alteração no arquivo

### Ator Primário

Usuário

### Fluxo de Eventos Principal

1. O usuário clica com o botão direito sobre uma sequência e escolhe a opção “Agrupar N’s” (**A1**)
2. O sistema apresenta a tela de confirmação de agrupamento (**DV1**)
3. O usuário clica em Ok (**A2**)
4. O sistema agrupa a sequência de caracteres N (**R1**)
5. O caso de uso é finalizado

## Fluxos Alternativos

**A1:** O usuário clica com o botão direito sobre uma sequência e escolhe a opção “Remover N’s”:

1. O sistema apresenta a tela de confirmação de remoção (**DV2**)
2. O usuário clica em Ok (**A2**)
3. O sistema remove a sequência de caracteres N (**R2**)
4. O caso de uso é finalizado.

**A2:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela de confirmação
2. O caso de uso é finalizado

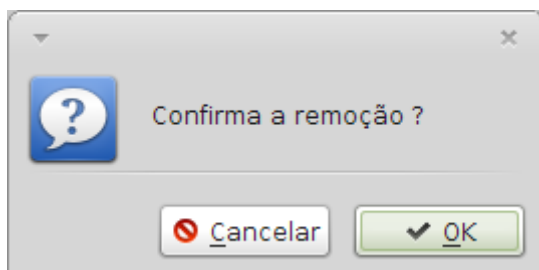
## Regras de Negócio

**R1:** O sistema agrupa em apenas um todas as sequências de caracteres “N” que estiverem lado a lado no arquivo (Ex.: TCATC**NNNNN**ATC para TCATC**N**ATC). Após o agrupamento deve-se atualizar a sequência no programa e no arquivo em que ela se encontra armazenada.

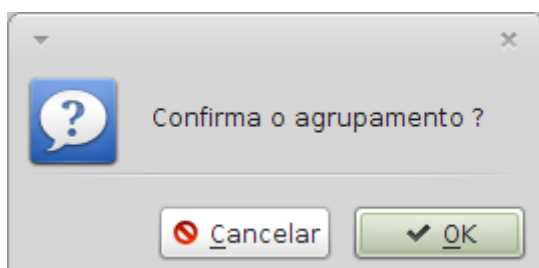
**R2:** O sistema realiza o mesmo processo da regra **R1**, porém ao invés de agrupar, remove todos os caracteres “N” (Ex.: TCATC**NNNNN**ATC para TCATCATC).

## Data View

**DV1 –** Tela de confirmação agrupamento



**DV2 –** Tela de confirmação exclusão





## UC08 – Editar sequência - Inserir

### Descrição

Este caso de uso descreve a inserção de parte de sequências

### Pré-condições

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

### Pós-condições

1. Ter inserido sequências informadas, replicando a alteração no arquivo

### Ator Primário

Usuário

### Fluxo de Eventos Principal

1. O usuário clica e marca uma posição na sequência (**R1**)
2. O usuário clica com o botão direito sobre a seleção e escolhe a opção “Inserir sequência”
3. O sistema apresenta a tela de entrada de texto (**DV1**)
4. O usuário digita a sequência e clica em Ok (**A1**)
5. O sistema insere a sequência informada (**R2**)
6. O caso de uso é finalizado

### Fluxos Alternativos

**A1:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela de confirmação sem realizar nenhuma ação.
2. O caso de uso é finalizado.

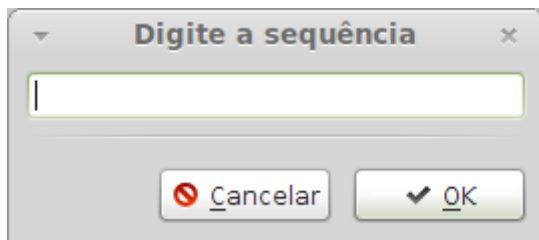
### Regras de Negócio

**R1:** O sistema deve marcar na tela com destaque a sequência selecionada pelo mouse, efetuada pelo clique.

**R2:** O sistema insere a sequência informada no campo de entrada no local onde a sequência havia sido marcada pela regra **R1** (Ex.: TCATCXXXXATC sendo XXXX a sequência digitada). Após a inserção deve-se atualizar a sequência no programa e no arquivo em que ela se encontra armazenada.

## Data View

### DV1 – Tela de entrada



### UC09 – Alinhar sequência

## Descrição

Este caso de uso descreve o alinhamento de uma sequência contig em relação à referência

## Pré-condições

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

## Pós-condições

1. Ter alinhado a sequência e atualizado sequencia consenso

## Ator Primário

Usuário

## Fluxo de Eventos Principal

1. O usuário clica com o botão direito em uma sequência e seleciona "Novo alinhamento"
2. O sistema apresenta a tela de alinhamento (**DV1**)

3. O usuário informa os dados na tela (**R1**)
4. O usuário clica em Alinhar (**A1**)
5. O sistema alinha a sequência contig selecionada (**R2**) (**A2**) (**E1**)
6. O sistema mostra a saída de alinhamento e relatório (**R3**) (**DV2**) (**DV3**)
7. O usuário clica em Fechar
8. O caso de uso é finalizado

### **Fluxos Alternativos**

**A1:** Botão “Fechar” pressionado:

1. Sistema fecha a tela de alinhamento.
2. O caso de uso é finalizado

**A2:** Botão “Cancelar” pressionado:

1. Sistema cancela processo atual de alinhamento que esteja em andamento.
2. O caso de uso é finalizado

### **Fluxos de Exceção**

**E1:** Erro no alinhamento:

1. O sistema trata o erro e informa no relatório posterior.
2. O caso de uso é finalizado

### **Regras de Negócio**

**R1:** O campo repetir informa quantas vezes o alinhamento deve ser feito naquela sequência. Os campos abaixo são parâmetros que serão utilizados nos algoritmos de alinhamento.

**R2:** O alinhamento deve ocorrer entre a(s) sequência selecionada (contigs) e a sequência de referência utilizando as bibliotecas do JContigSort, que irá identificar uma mediana de alinhamento. Com este parâmetro de mediana o sistema deve realizar o alinhamento local pela biblioteca Jaligner, que irá retornar o posicionamento do contig alinhado.

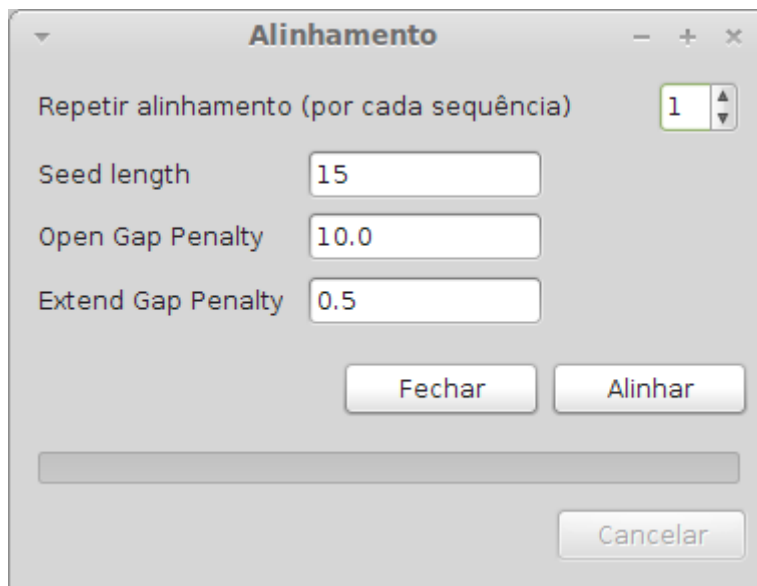
Após possuir os dados de alinhamento o sistema deve posicionar a sequência alinhada no local correto na tela, deve também posicioná-la no mapa e atualizar a sequência consenso.

**R3:** Os valores de retorno e o resultado do alinhamento devem ser mostrados na saída de alinhamento, conforme a **DV2**.

O relatório deve informar qual a situação de alinhamento das sequências alinhadas.

## Data View

### DV1 – Tela alinhamento



**Alinhamento**

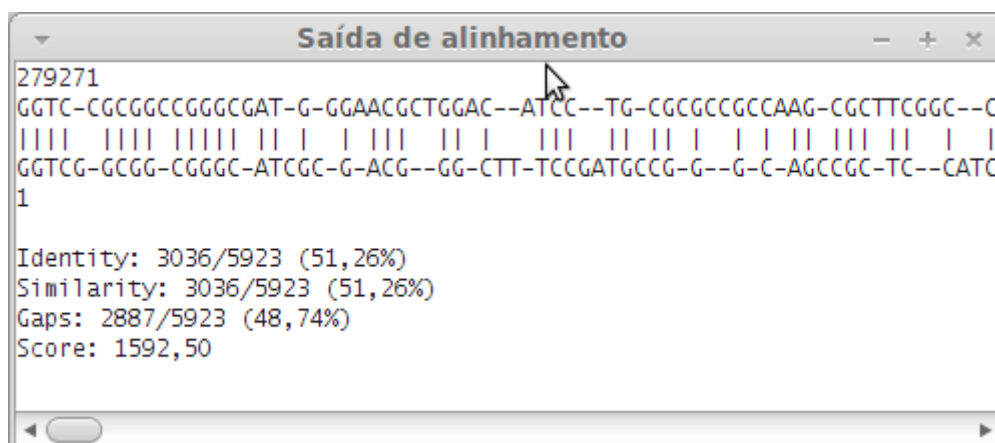
Repetir alinhamento (por cada sequência)

Seed length

Open Gap Penalty

Extend Gap Penalty

### DV2 – Saída alinhamento



**Saída de alinhamento**

279271

GGTC-CGCGGCCGGGCGAT-G-GGAACGCTGGAC--ATCC--TG-CGCGCCGCCAAG-CGCTTCGGC--C

|||| | ||| | ||| | | | | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

GGTCG-GCGG-CGGGC-ATCGC-G-ACG--GG-CTT-TCCGATGCCG-G--G-C-AGCCGC-TC--CATC

1

Identity: 3036/5923 (51,26%)

Similarity: 3036/5923 (51,26%)

Gaps: 2887/5923 (48,74%)

Score: 1592,50

**DV3 – Relatório****UC10 – Visualizar alinhamento****Descrição**

Este caso de uso descreve a visualização dos alinhamentos na interface do programa

**Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

**Pós-condições**

1. Ter visualizado os alinhamentos corretamente

**Ator Primário**

Usuário



### **Pós-condições**

1. Ter exportado a sequência para o formato desejado (mapa, sequência consenso ou sequências)

### **Ator Primário**

Usuário

### **Fluxo de Eventos Principal**

1. O usuário seleciona uma opção no menu "Exportar" (**DV1**)
2. O sistema apresenta a tela de salvar arquivo (**DV2**)
3. O usuário seleciona um nome para o arquivo
4. O usuário clica no botão "Ok" (**A1**)
5. O sistema exporta o conteúdo selecionado (**R1**) (**E1**)
6. O caso de uso é finalizado

### **Fluxos Alternativos**

**A1:** Botão "Cancelar" pressionado:

1. Sistema fecha a tela para salvar arquivo
2. O caso de uso é finalizado

### **Fluxos de Exceção**

**E1.** Memória insuficiente para gerar imagem:

1. O sistema apresenta a mensagem de erro (**DV3**)
2. O caso de uso é reiniciado

### **Regras de Negócio**

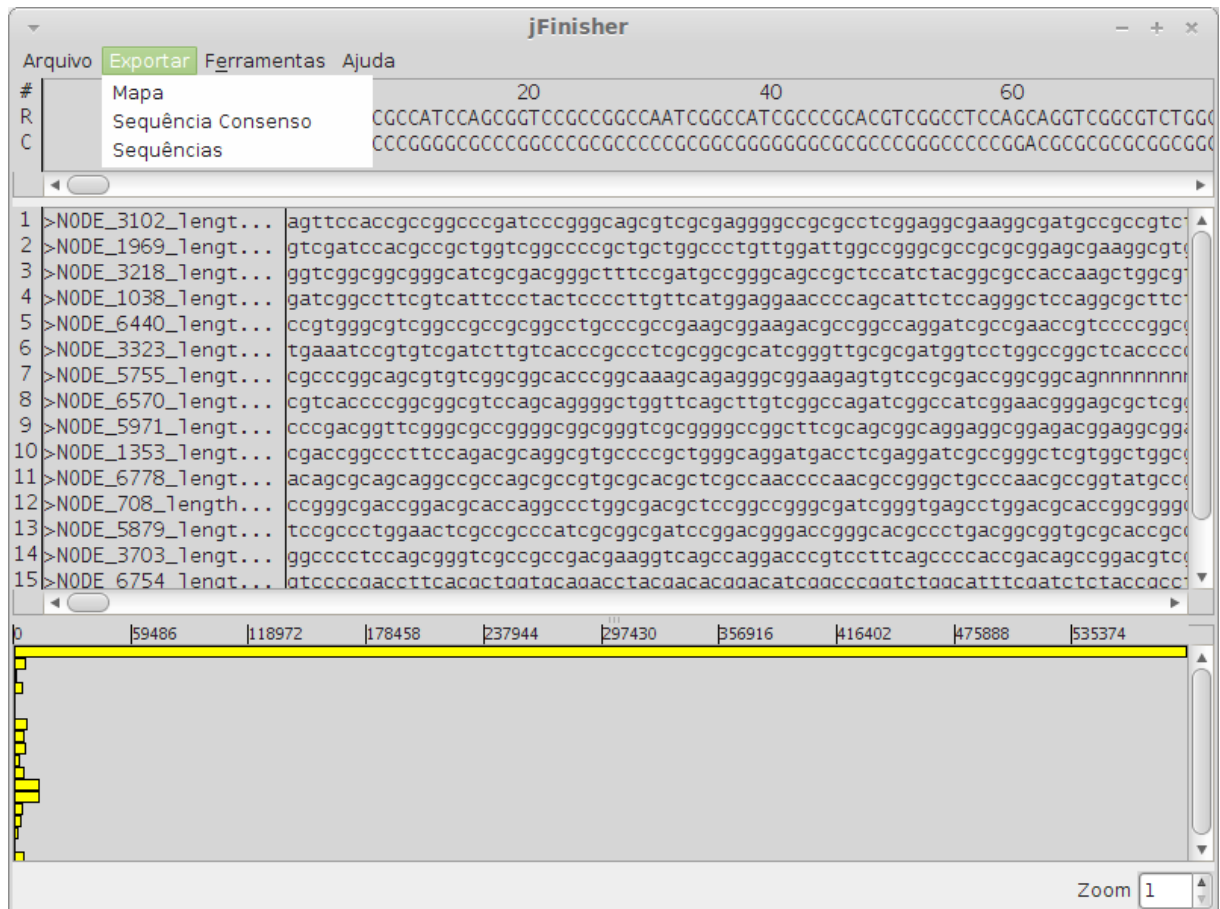
**R1:** Mapa: Sistema exporta imagem do mapa de sequências no formato PNG (.png)

Sequência consenso: Sistema exportar a sequência consenso no formato FASTA (.fasta)

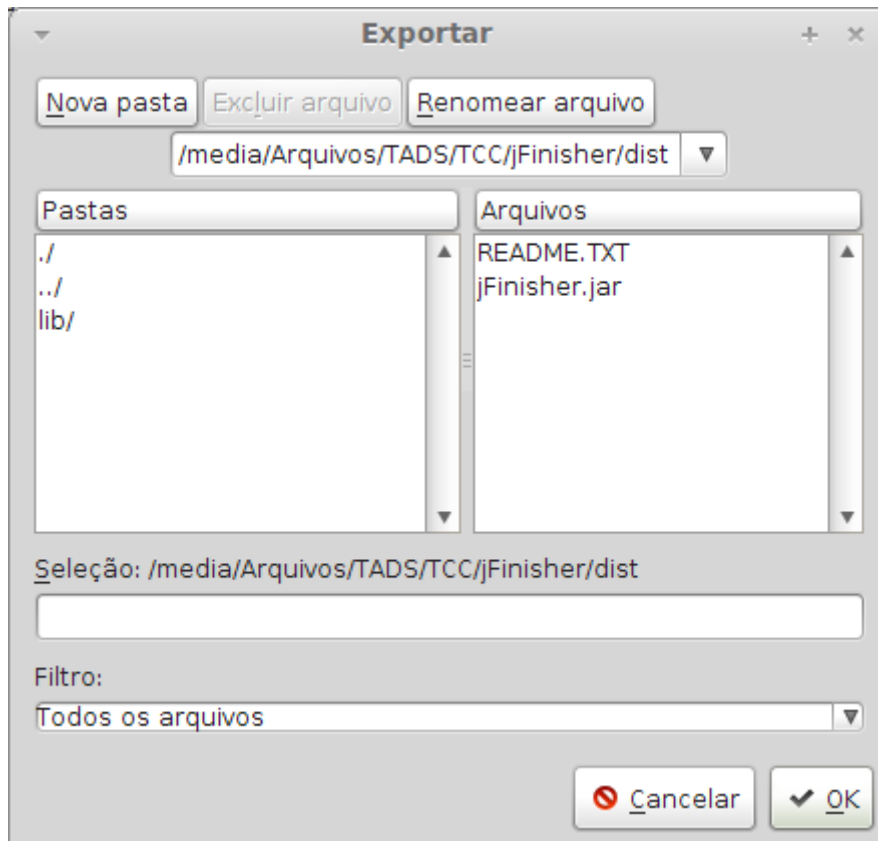
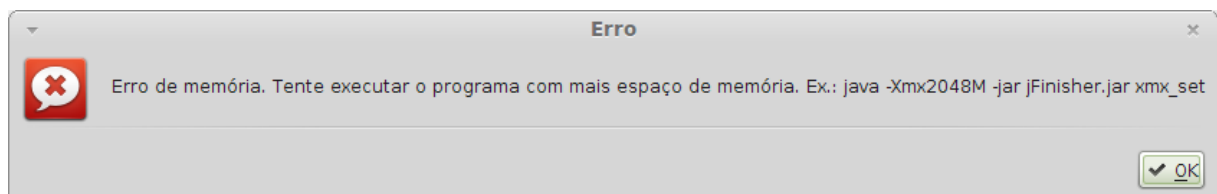
Sequências: Sistema exporta as sequências abertas no formato MULTI FASTA (.ffn)

### **Data View**

## DV1 - Opções para exportar





**DV2 - Tela para salvar arquivos****DV3 - Tela de erro****UC12 – Agrupar sequências****Descrição**

Este caso de uso descreve como será realizado o agrupamento de contigs

**Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter um ou mais arquivos de contig abertos

### Pós-condições

1. Ter agrupado as sequência conforme opções escolhidas

### Ator Primário

Usuário

### Fluxo de Eventos Principal

1. O usuário seleciona a opção “Ferramentas, Unir sequências” no menu
2. O sistema apresenta a tela de unificar sequências (**DV1**)
3. O usuário marca “Unir todas as sequências” e “Criar nova sequência” e clica em “Ok” (**A1**) (**A2**) (**A3**)
4. O sistema apresenta a tela de salvar arquivo (**DV2**)
5. O usuário digita o nome do arquivo e clica no botão “Ok” (**A4**)
6. O sistema unifica as sequências conforme parâmetros (**R1**)
7. O caso de uso é finalizado

### Fluxos Alternativos

**A1:** Usuário marca “Unir sequências específicas”:

1. O caso de uso volta para o fluxo principal com opção marcada

**A2:** Usuário marca “Substituir primeira sequência”:

1. O caso de uso volta para o fluxo principal com opção marcada e ignora passo 4 e 5

**A3:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela para unificar sequências
2. O caso de uso é finalizado

**A4:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela para salvar arquivo
2. O caso de uso é finalizado

### Regras de Negócio

**R1:** Unir todas as sequências: Sistema une todas as sequências em uma só

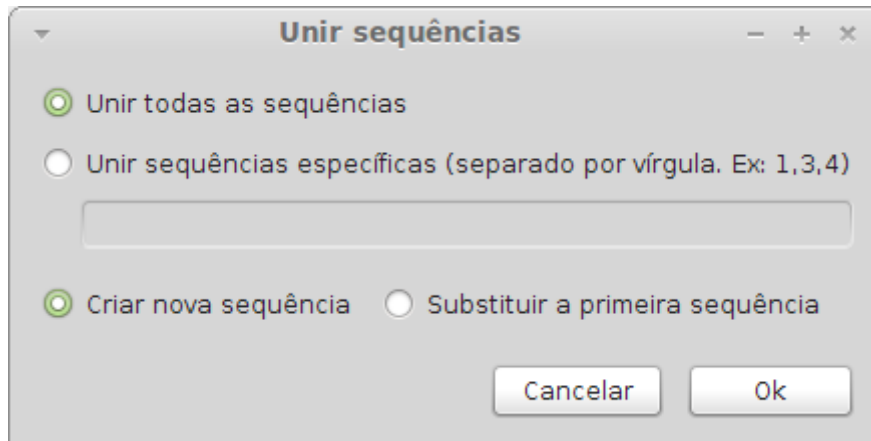
Unir sequências específicas: Sistema apenas une as sequências informadas pelo usuário no campo de entrada

Criar nova sequência: Sistema gera novo arquivo com sequências unidas

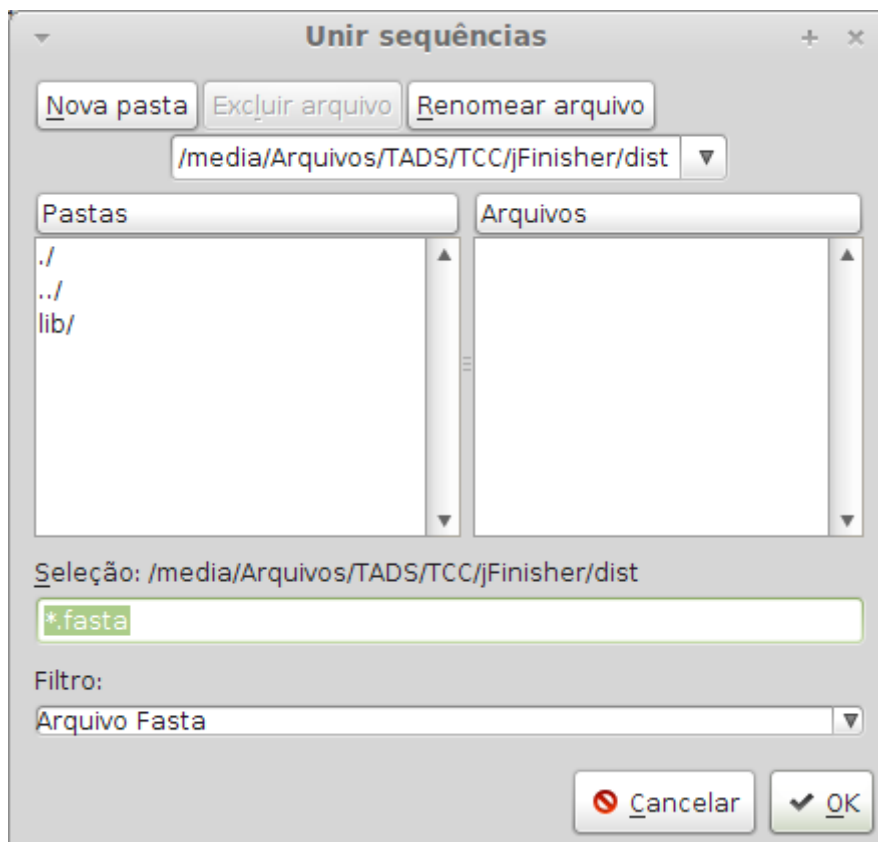
Substituir primeira sequência: Sistema gera sequência unida e sobrepõe a primeira sequência do conjunto unificado

## Data View

### DV1 – Tela para unificar sequências



### DV2 – Salvar arquivo



## UC13 – Novo projeto

### Descrição

Este caso de uso descreve a criação de um novo projeto

### Pré-condições

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter (não obrigatório) um ou mais arquivos de contig abertos

### Pós-condições

1. Ter iniciado um novo projeto e fechado projeto previamente aberto

### Ator Primário

Usuário

### Fluxo de Eventos Principal

1. O usuário clica em “Arquivo, Novo Projeto” no menu
2. O sistema apresenta o aviso (**DV1**)
3. O usuário clica em “Sim” (**A1**)
4. O sistema inicia um novo projeto (**R1**)
5. O caso de uso é finalizado

### Fluxos Alternativos

**A1:** Botão “Não” pressionado:

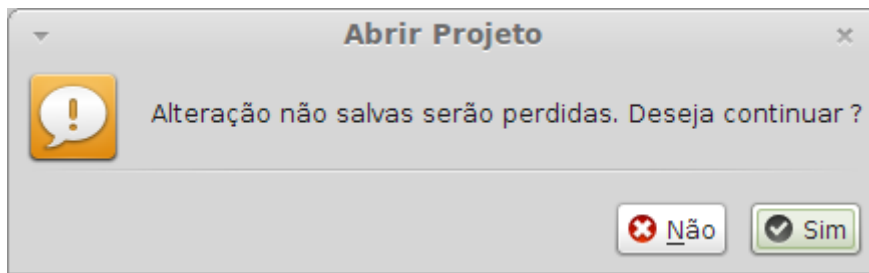
1. Sistema fecha a tela de aviso sem executar nenhuma ação
2. O caso de uso é reiniciado

### Regras de Negócio

**R1:** Sistema retira todos os arquivos abertos da tela e retoma sistema na posição inicial.

## Data View

### DV1 – Aviso



### UC14 – Abrir projeto

## Descrição

Este caso de uso descreve a abertura de um projeto

## Pré-condições

1. O sistema deve estar iniciado

## Pós-condições

1. Ter aberto o projeto e incluído os arquivos referentes

## Ator Primário

Usuário

## Fluxo de Eventos Principal

1. O usuário clica em “Arquivo, Abrir Projeto” no menu
2. O sistema apresenta a tela para abrir projeto (**DV1**)
3. O usuário seleciona um arquivo e clica em “Ok” (**A1**) (**E1**)
4. O sistema abre o projeto (**R1**)
5. O caso de uso é finalizado

## Fluxos Alternativos

**A1:** Botão “Cancelar” pressionado:

1. Sistema fecha a tela para abrir projeto
2. O caso de uso é finalizado

## Fluxos de Exceção

### E1. Arquivo não identificado:

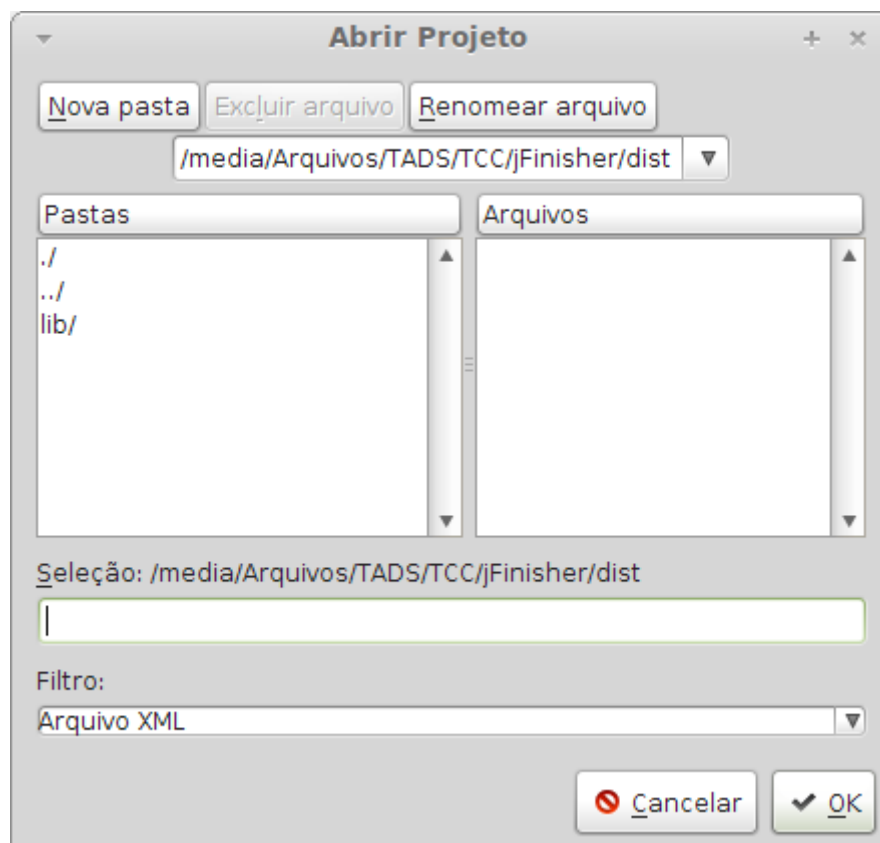
1. O apresenta a mensagem de erro (**DV2**)
2. O caso de uso é reiniciado

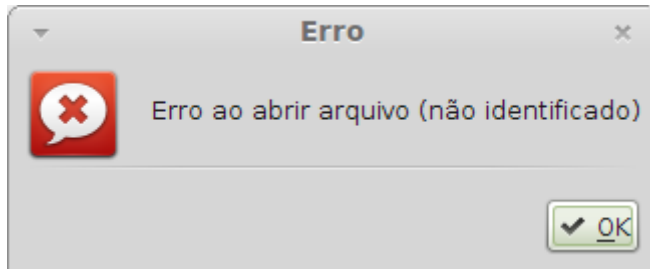
## Regras de Negócio

**R1:** Sistema abre sequências informadas no arquivo de projeto. Caso apresente alguma incompatibilidade ou não encontre os arquivo de origem, informa ao usuário as sequências específicas que não foram encontradas e pede para localizá-las.

## Data View

### DV1 – Abrir projeto



**DV2 – Tela de erro****UC15 – Salvar projeto****Descrição**

Este caso de uso descreve como salvar um projeto

**Pré-condições**

1. O sistema deve estar iniciado
2. Ter o arquivo de referência aberto
3. Ter (não obrigatório) um ou mais arquivos de contig abertos

**Pós-condições**

1. Ter salvo o projeto e incluído os arquivos referentes

**Ator Primário**

Usuário

**Fluxo de Eventos Principal**

1. O usuário clica em “Arquivo, Salvar Projeto” no menu
2. O sistema apresenta a tela para salvar projeto (**DV1**)
3. O usuário informa um novo arquivo e clica em “Ok” (**A1**)
4. O sistema salva o projeto (**R1**)
5. O caso de uso é finalizado

**Fluxos Alternativos**

**A1:** Botão “Cancelar” pressionado:

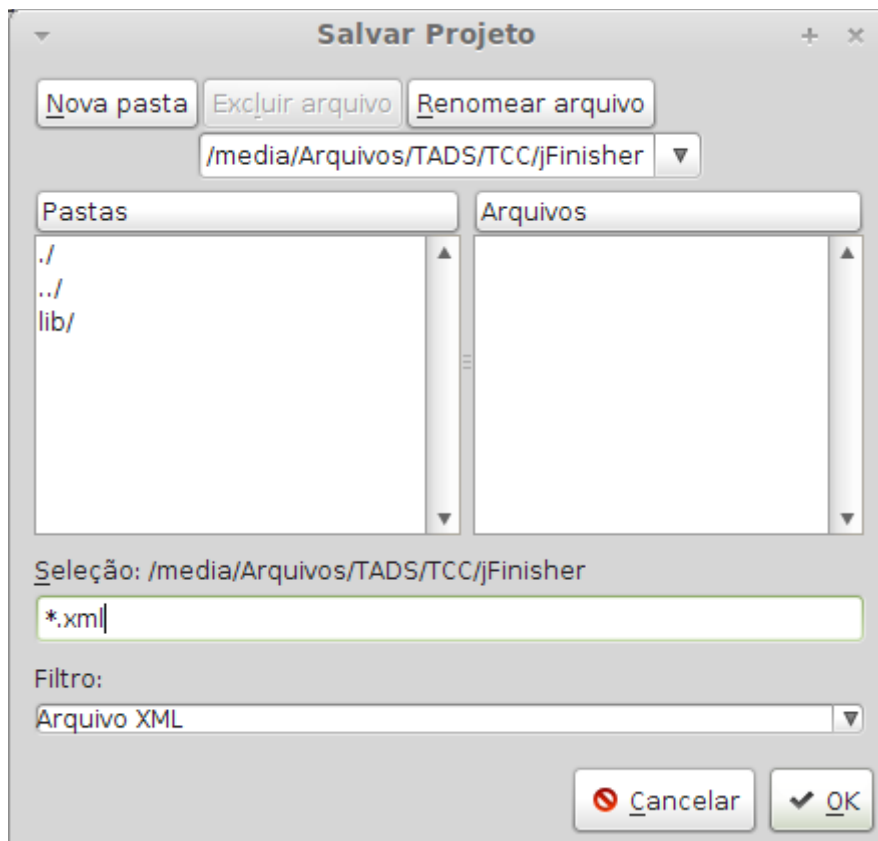
1. Sistema fecha a tela para salvar projeto
2. O caso de uso é finalizado

## Regras de Negócio

**R1:** Sistema deve salvar o projeto em um arquivo XML (.xml) armazenando as informações necessárias sobre a sequência e seus alinhamento para poder realizar a abertura posterior dos dados, tendo todas as informações disponíveis. Não é necessário armazenar as sequências no arquivo de projeto, somente a referência para os arquivos originais.

## Data View

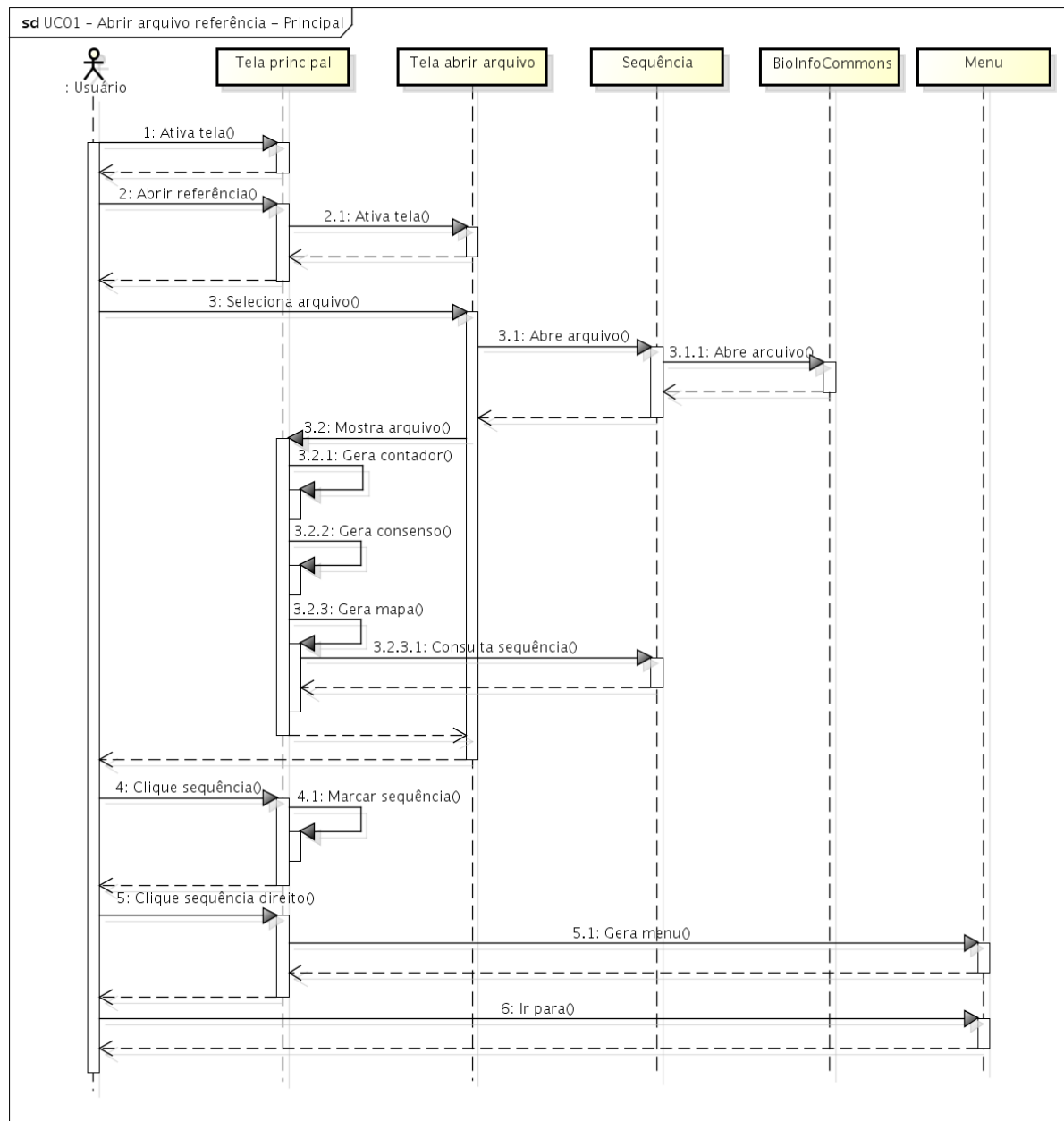
### DV1 – Salvar projeto





## 4 DIAGRAMAS DE SEQUÊNCIA

### UC01 – Abrir arquivo referência



powered by astah®

Figura 30 – Diagrama de Sequência – Abrir arquivo referência - Principal

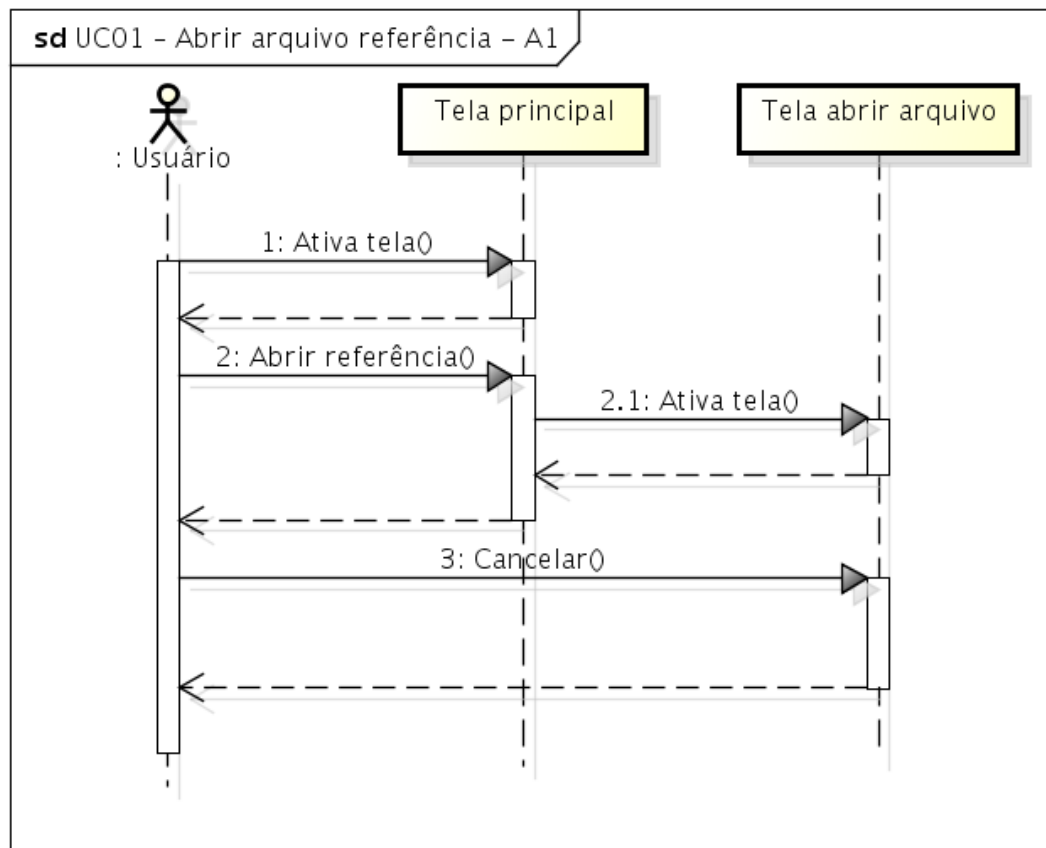
powered by astah<sup>®</sup>

Figura 31 – Diagrama de Sequência – Abrir arquivo referência – A1

## UC02 – Abrir-Adicionar arquivo(s) de sequência-contigs

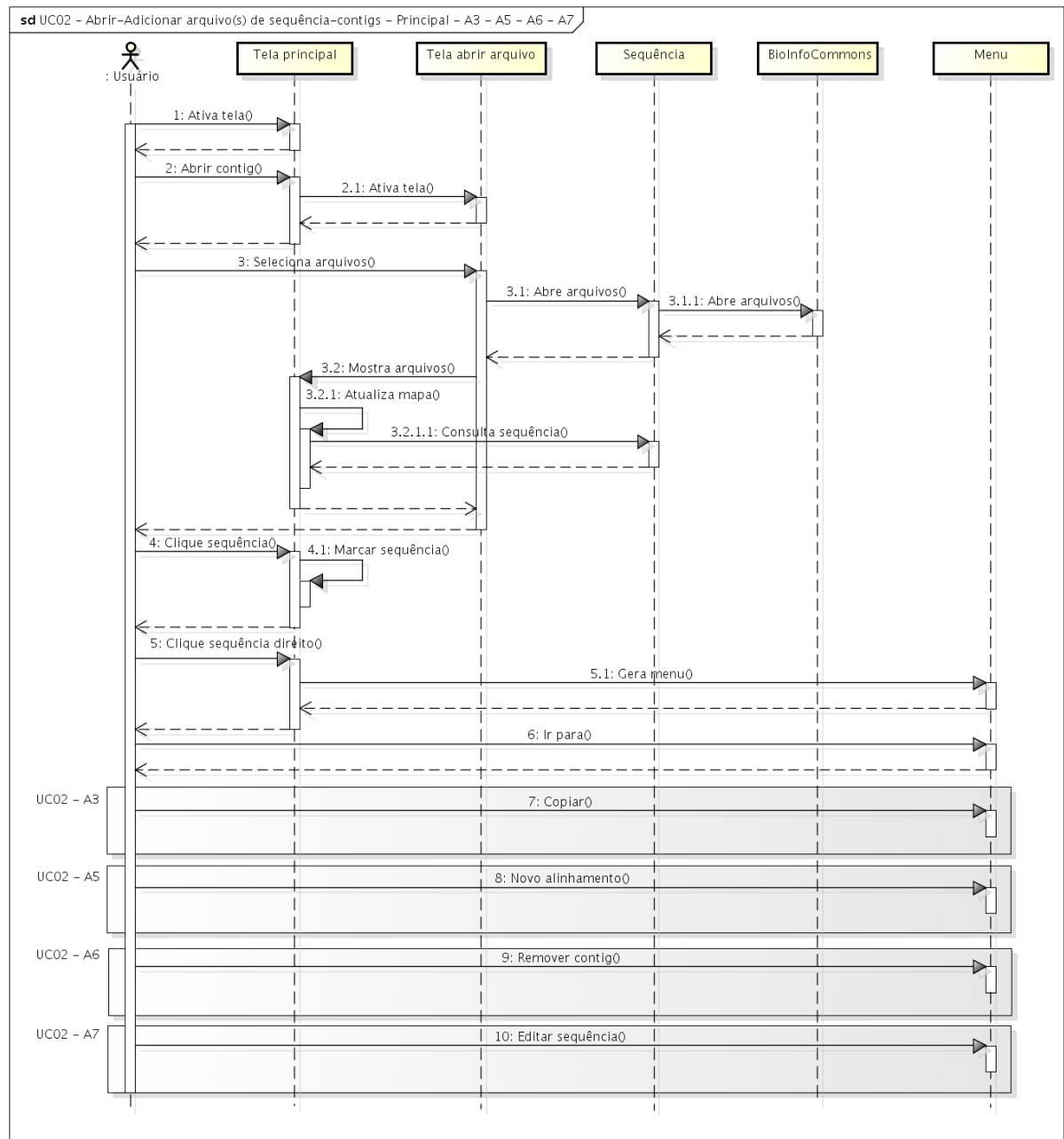
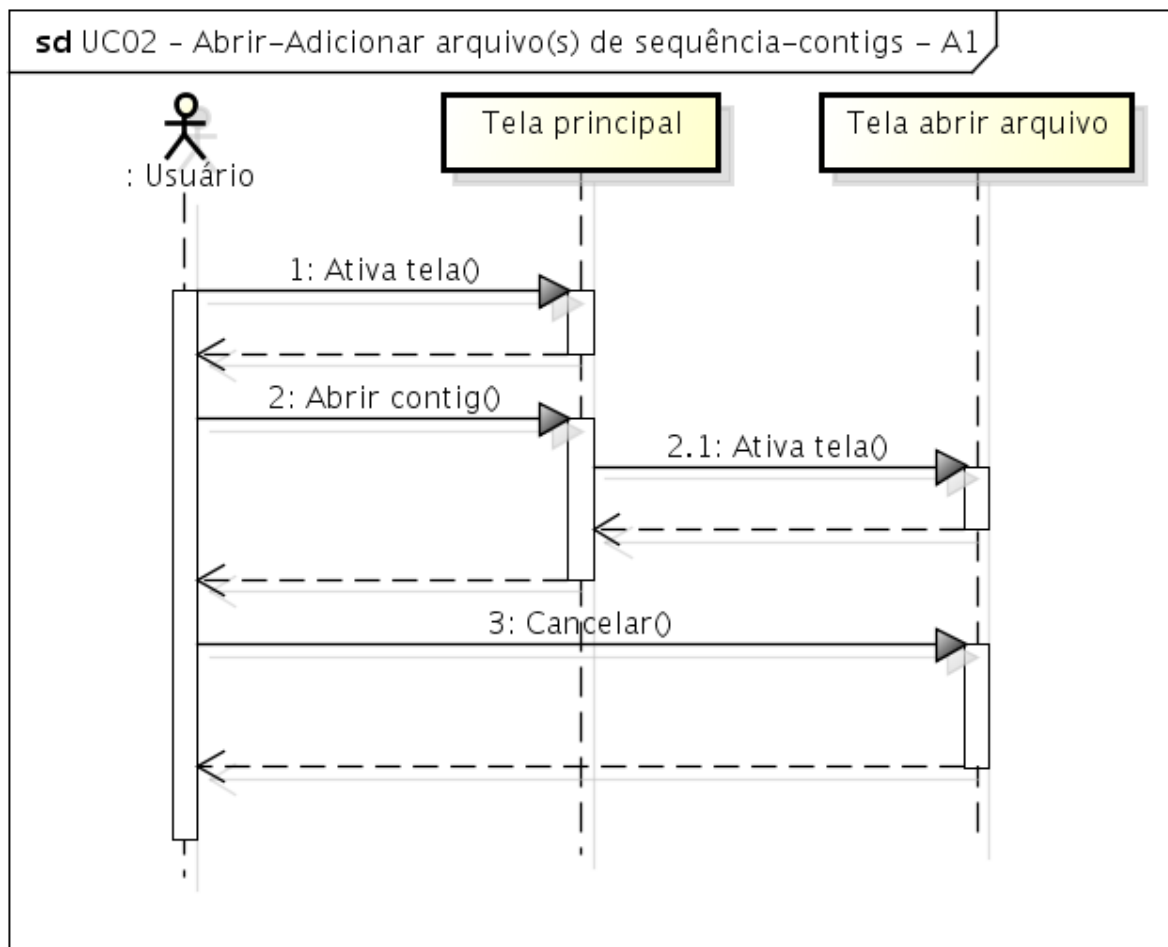
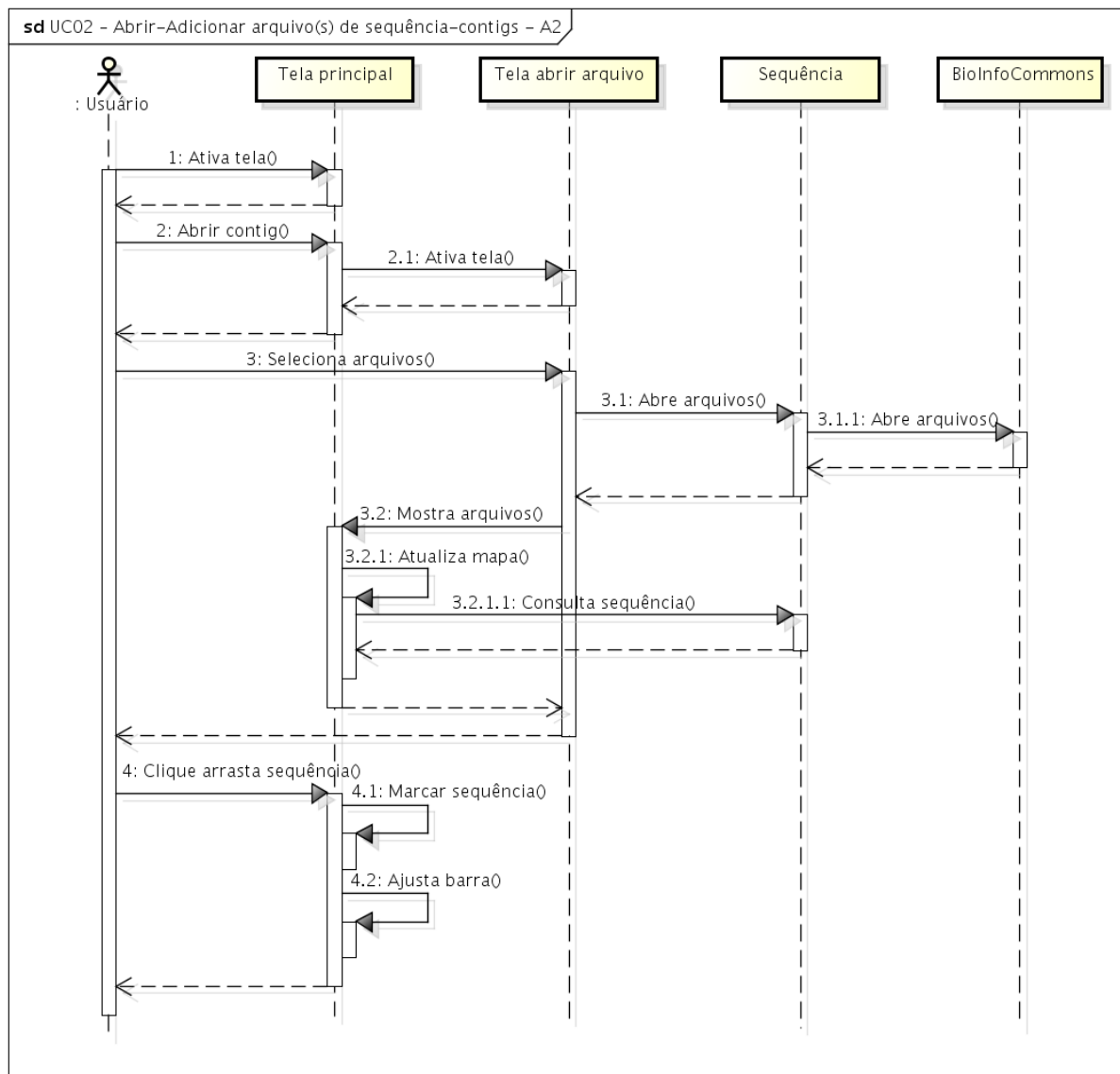


Figura 32 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência-contigs – Principal – A3 - A5 - A6 - A7



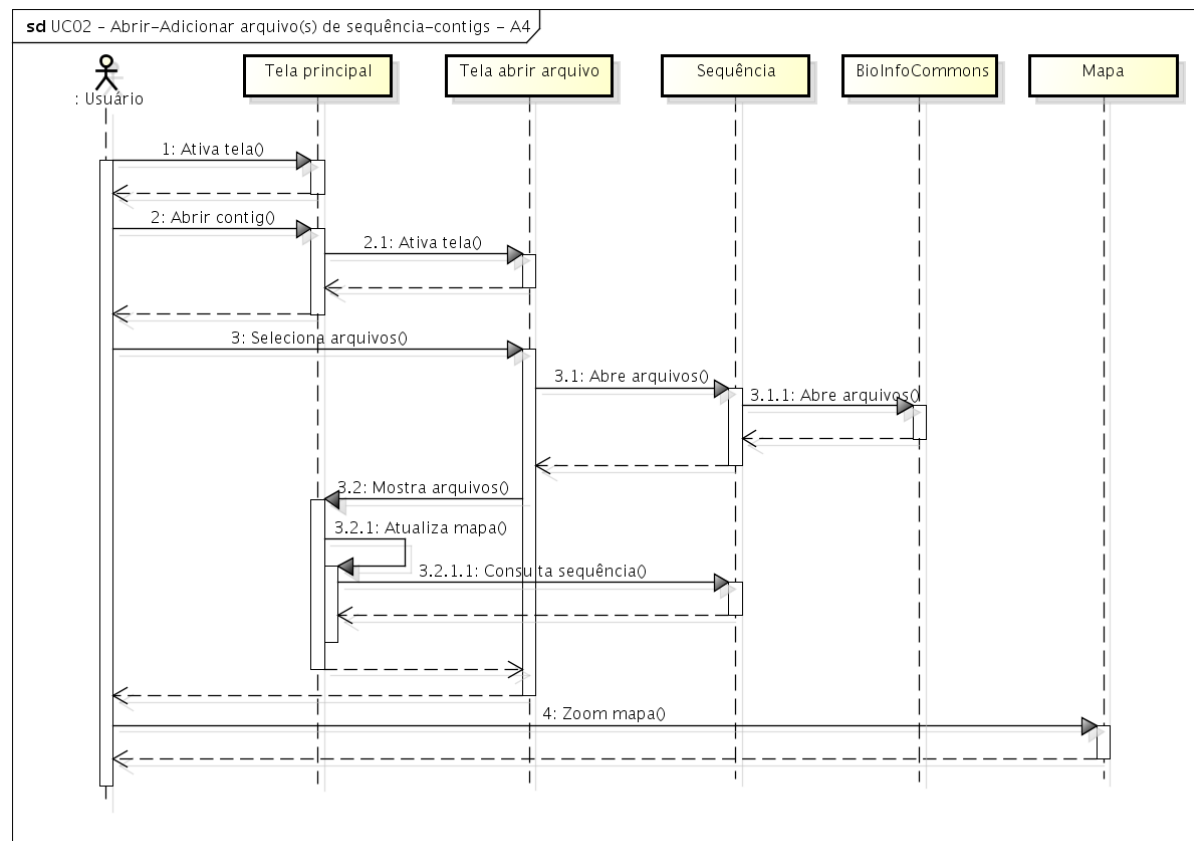
powered by astah\*

Figura 33 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência-contigs – A1



powered by astah®

Figura 34 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência-contigs – A2



powered by astah

Figura 35 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência-contigs – A4

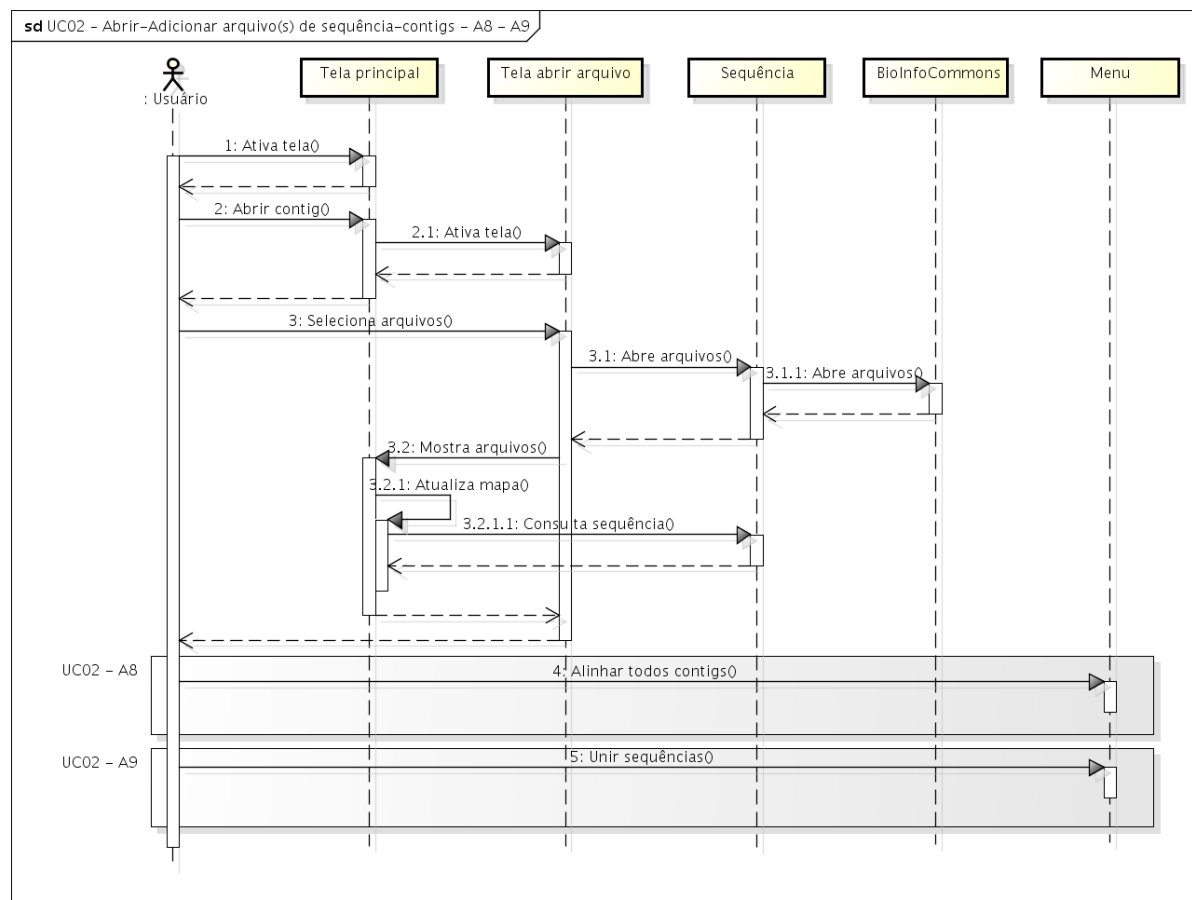
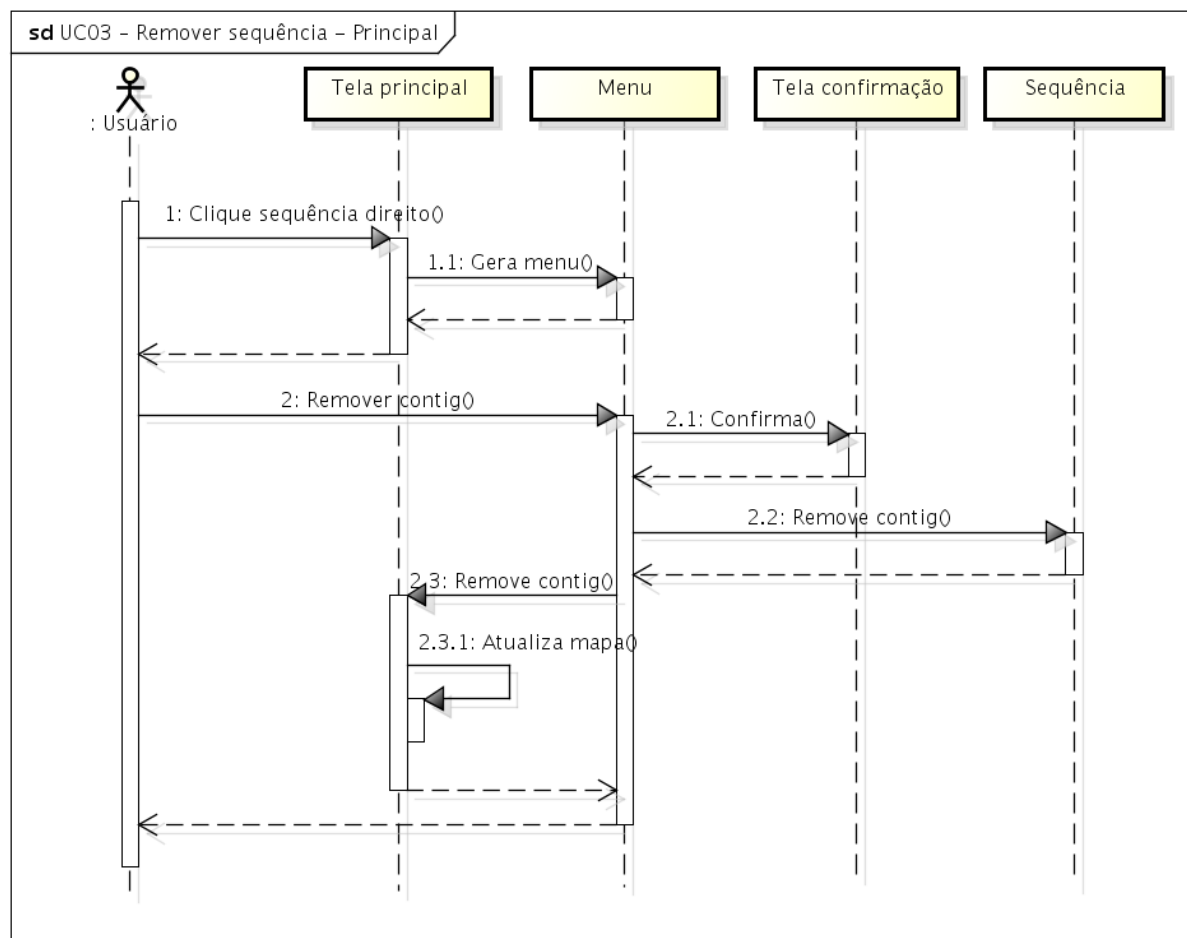


Figura 36 – Diagrama de Sequência – Abrir-Adicionar arquivo(s) de sequência-contigs – A8 – A9

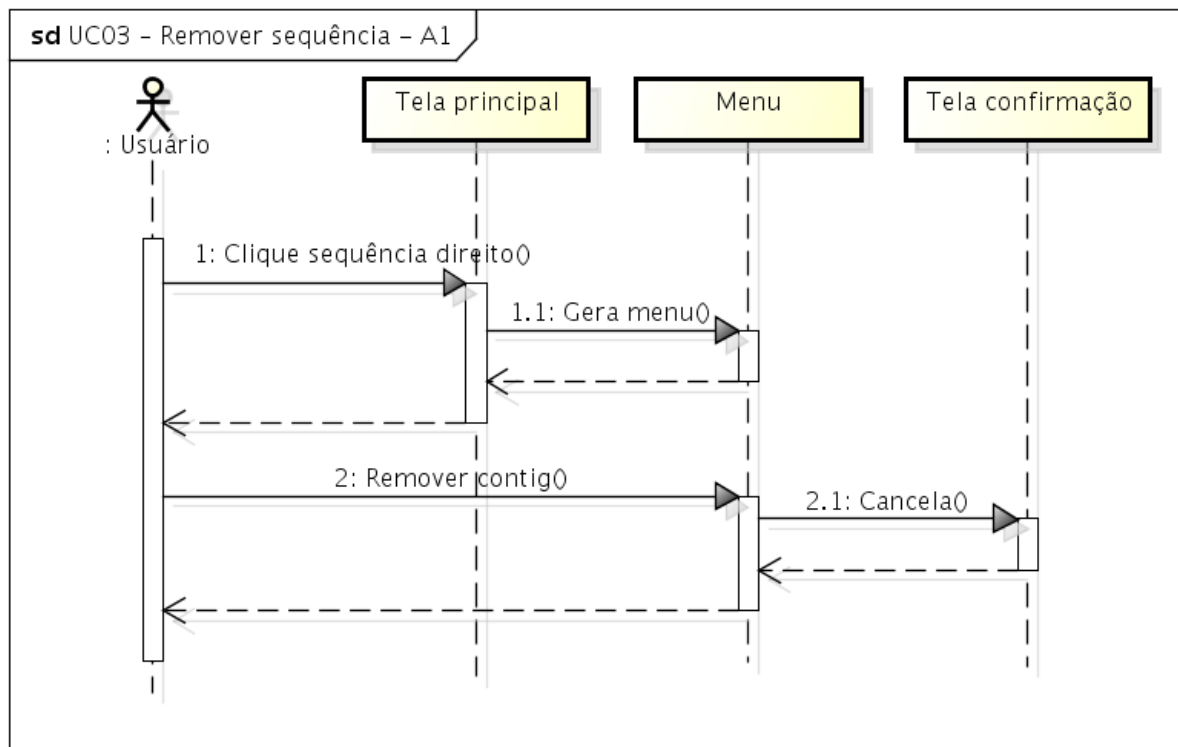
## UC03 – Remover sequência



powered by astah®

Figura 37 – Diagrama de Sequência – Remover sequência - Principal

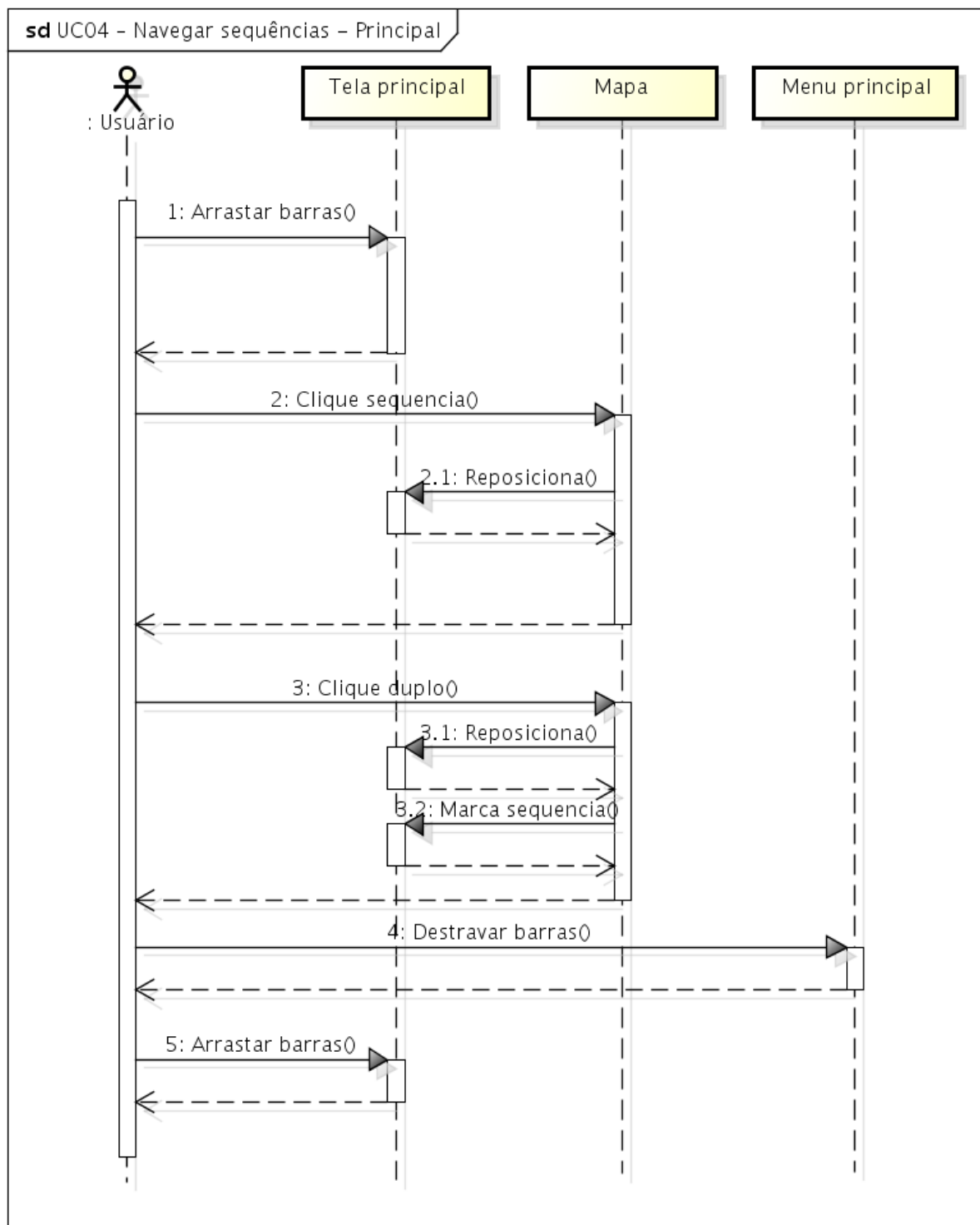




powered by astah®

Figura 38 – Diagrama de Sequência – Remover sequência – A1

## UC04 – Navegar sequências-mapa



powered by astah®

Figura 39 – Diagrama de Sequência – Navegar sequências-mapa - Principal

## UC05 – Localizar-Ir para posição

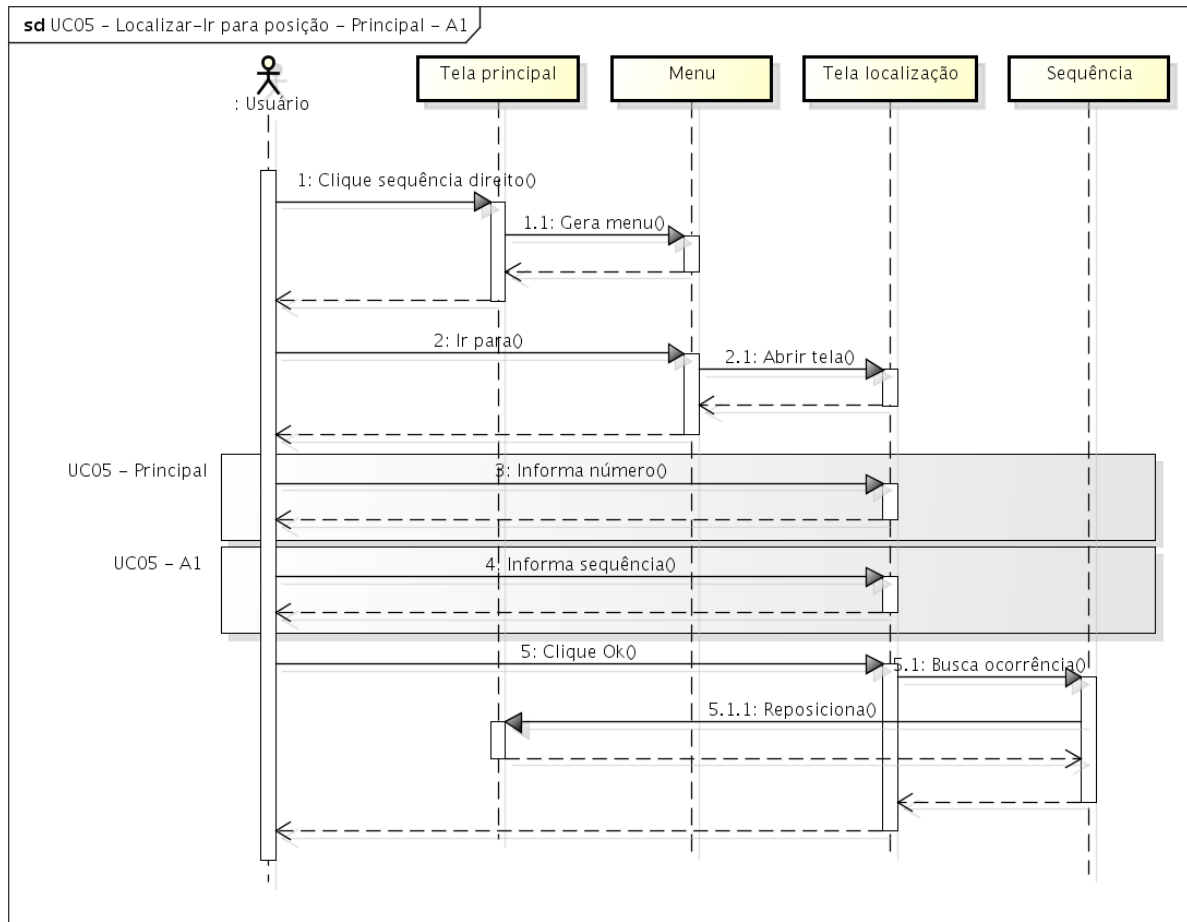
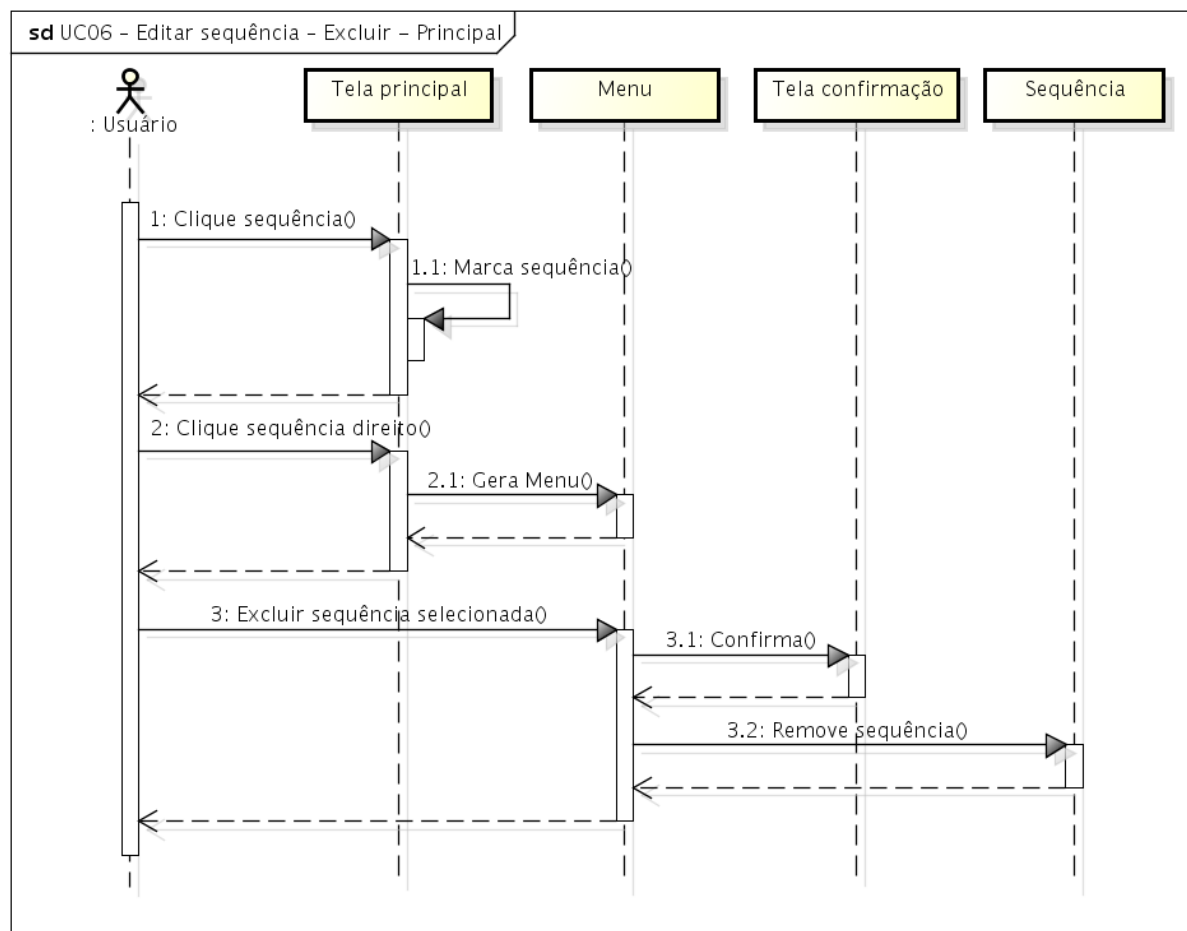


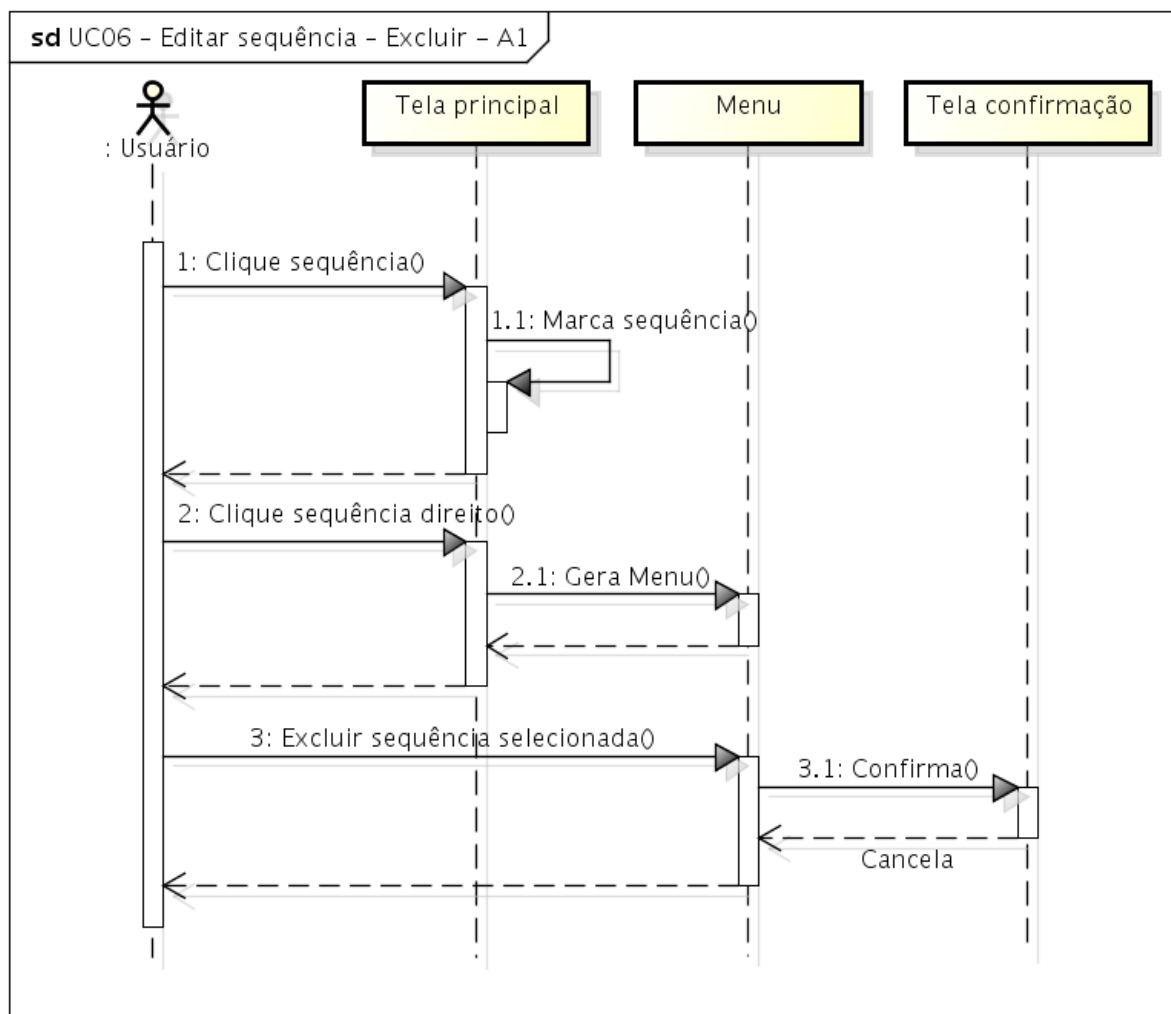
Figura 40 – Diagrama de Sequência – Localizar – Ir para posição – Principal – A1

## UC06 – Editar sequência – Excluir



powered by astah®

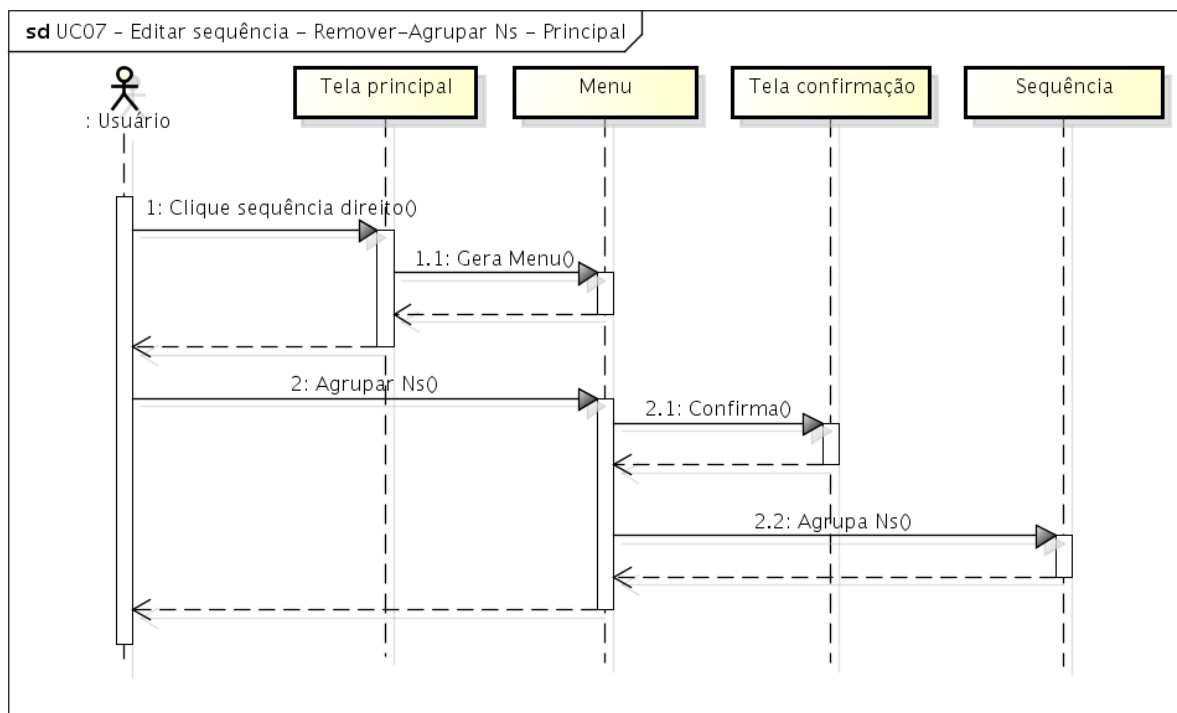
Figura 41 – Diagrama de Sequência – Editar sequência – Excluir - Principal



powered by astah®

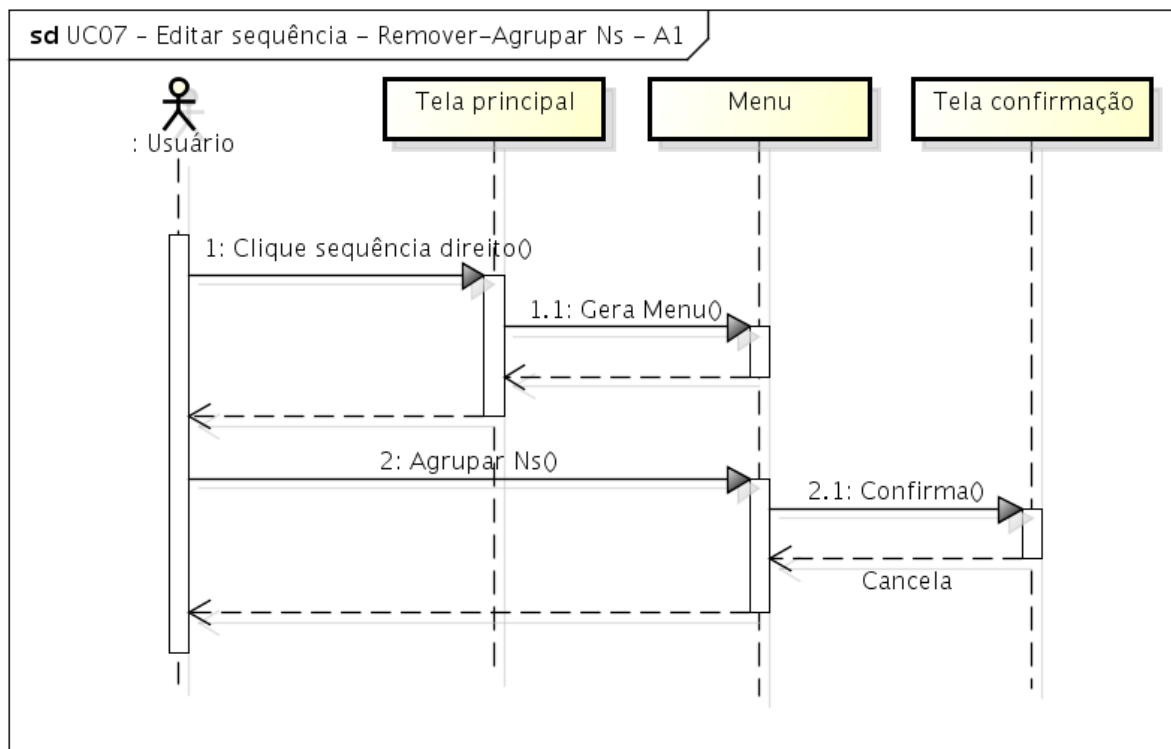
Figura 42 – Diagrama de Sequência – Editar sequência – Excluir – A1

## UC07 – Editar sequência - Remover-Agrupar Ns



powered by astah®

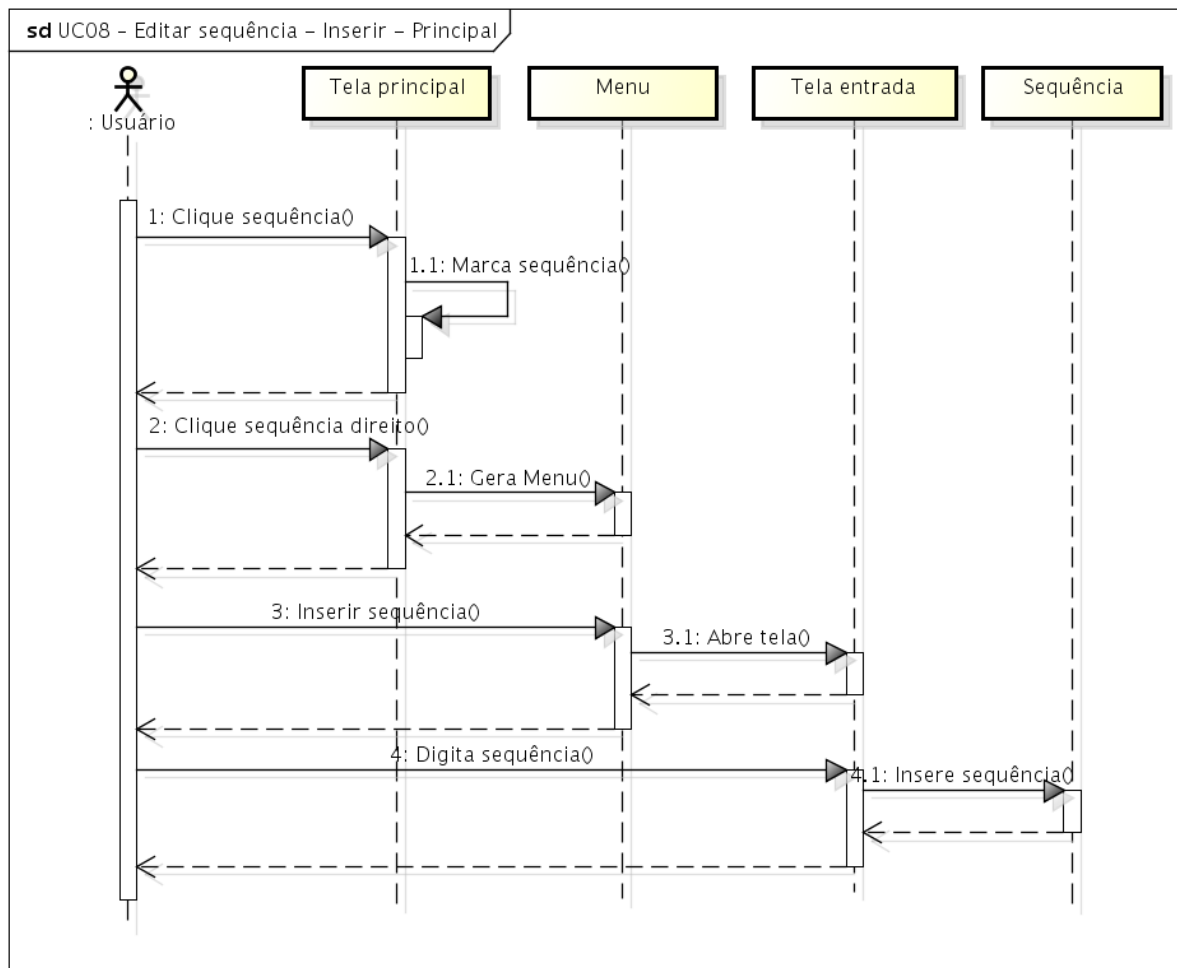
Figura 43 – Diagrama de Sequência – Editar sequência – Remover-Agrupar Ns – Principal



powered by astah®

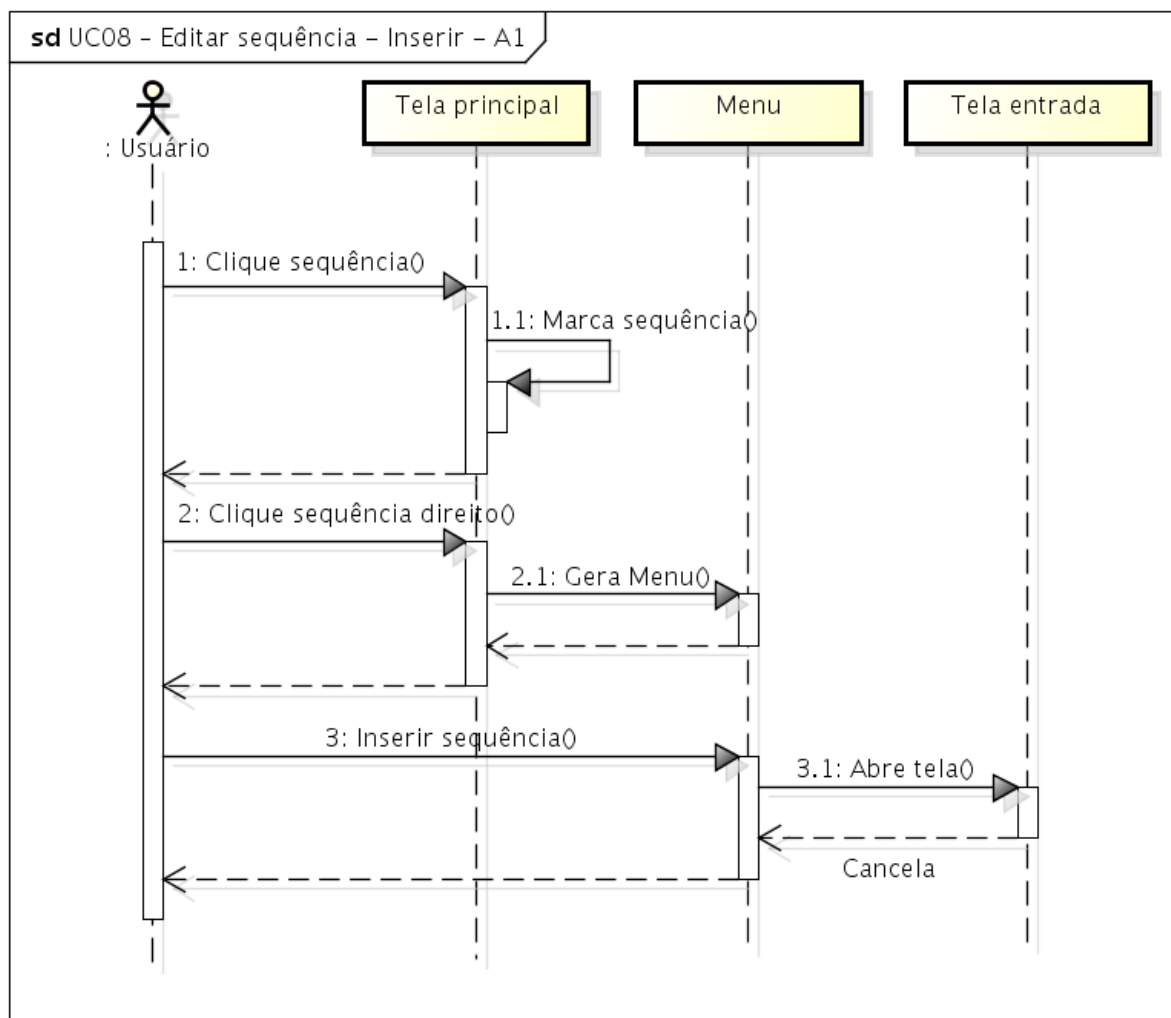
Figura 44 – Diagrama de Sequência – Editar sequência – Remover-Agrupar Ns – A1

## UC08 – Editar sequência – Inserir



powered by astah®

Figura 45 – Diagrama de Sequência – Editar sequência – Inserir – Principal

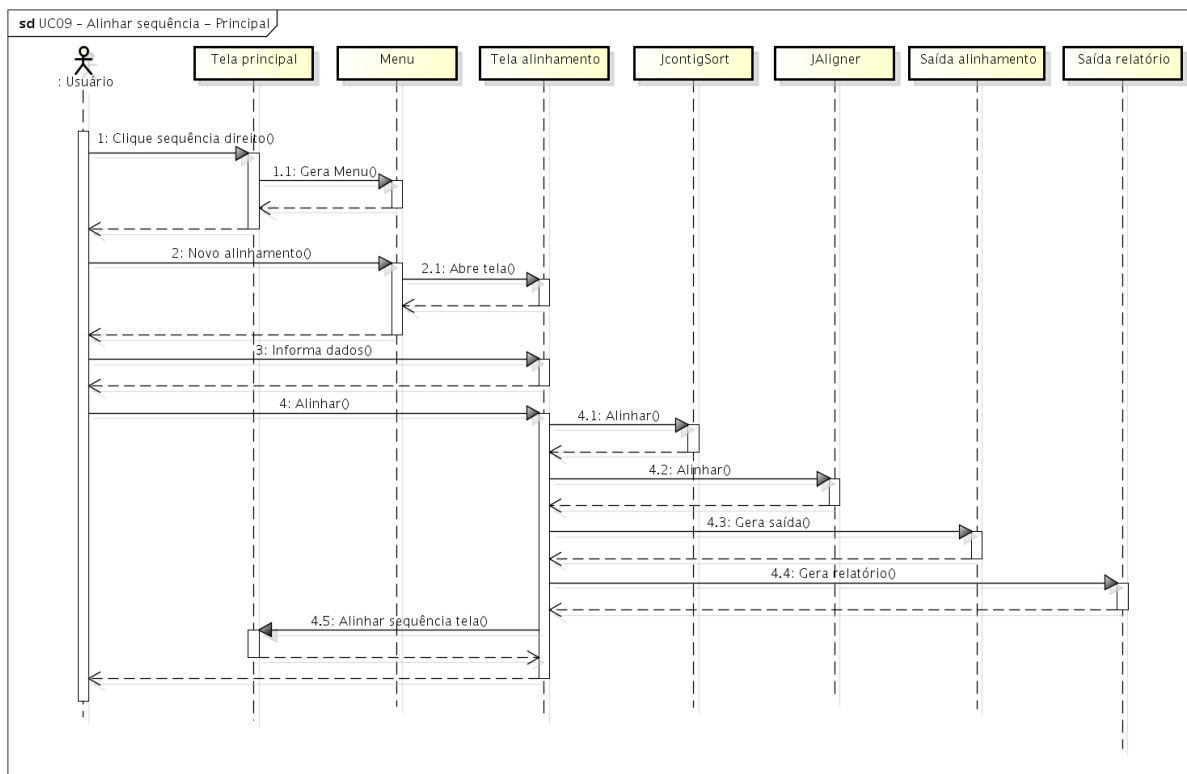


powered by astah®

Figura 46 – Diagrama de Sequência – Editar sequência – Inserir – A1



## UC09 – Alinhar sequência



powered by astah

Figura 47 – Diagrama de Sequência – Alinhar sequência – Principal

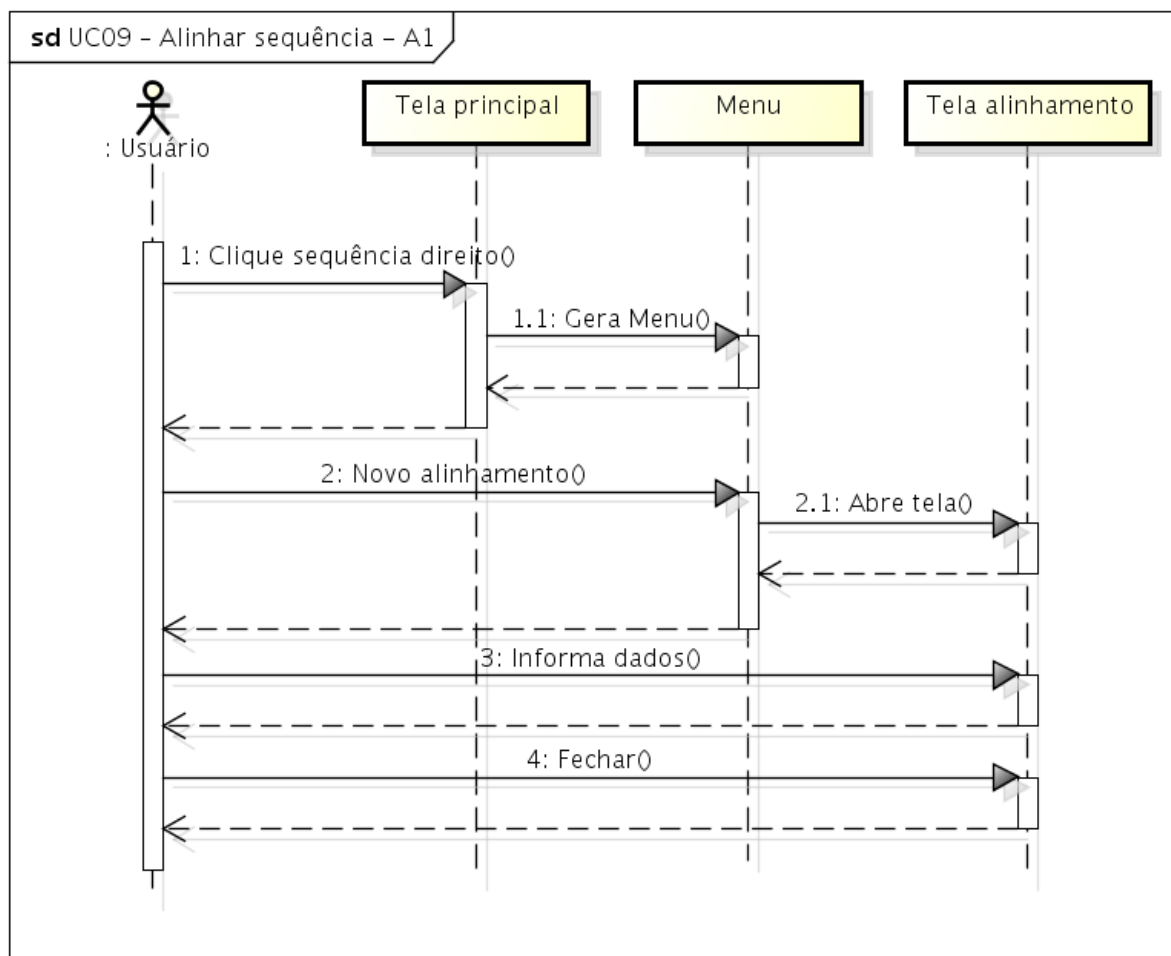


Figura 48 – Diagrama de Sequência – Alinhar sequência – A1

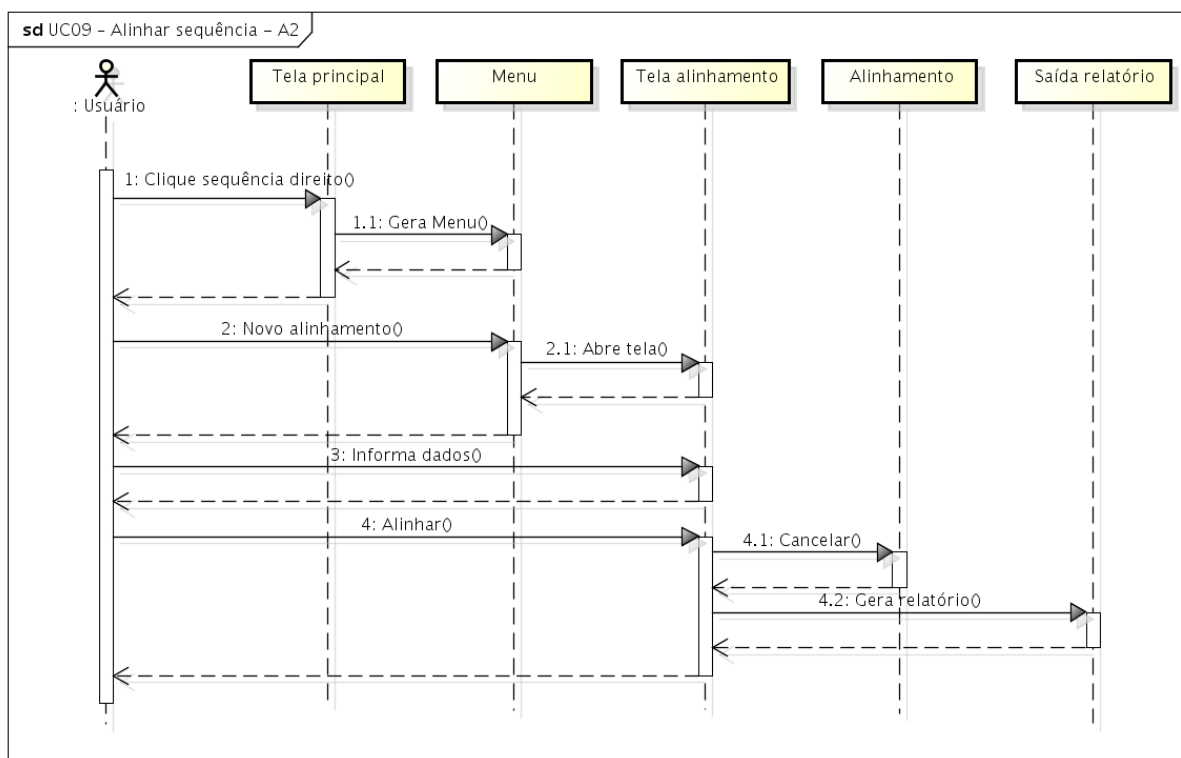


Figura 49 – Diagrama de Sequência – Alinhar sequência – A2

## UC10 – Visualizar alinhamento

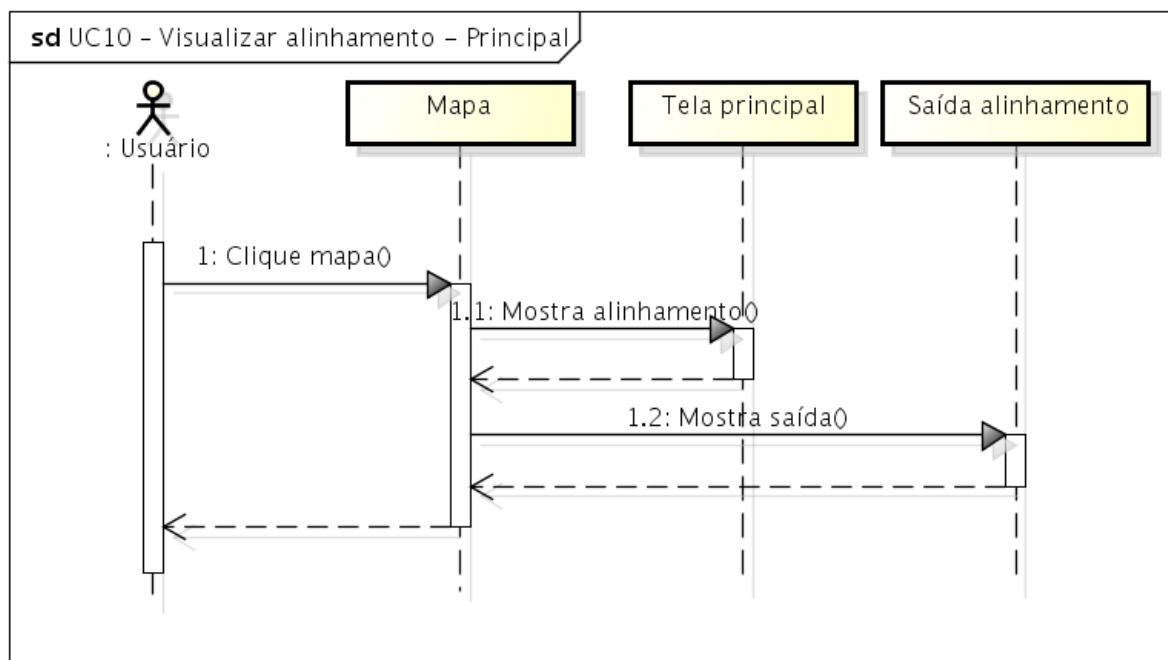
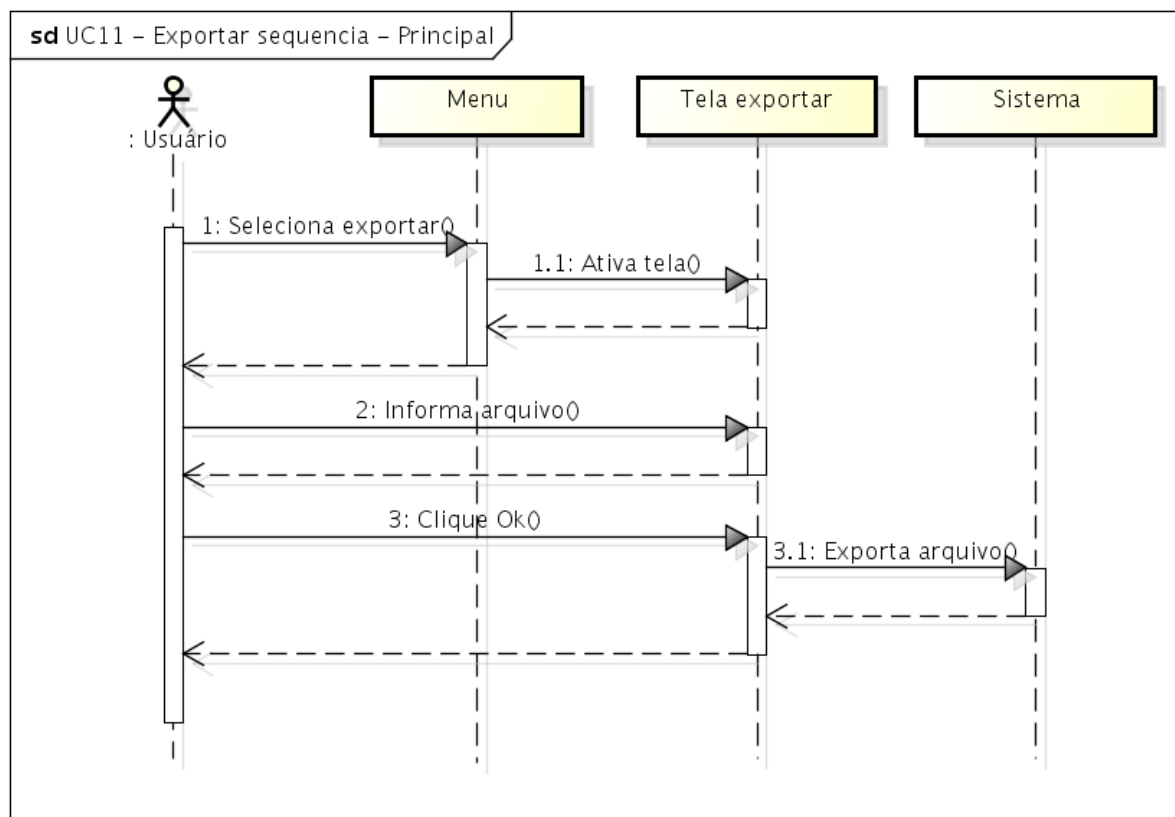


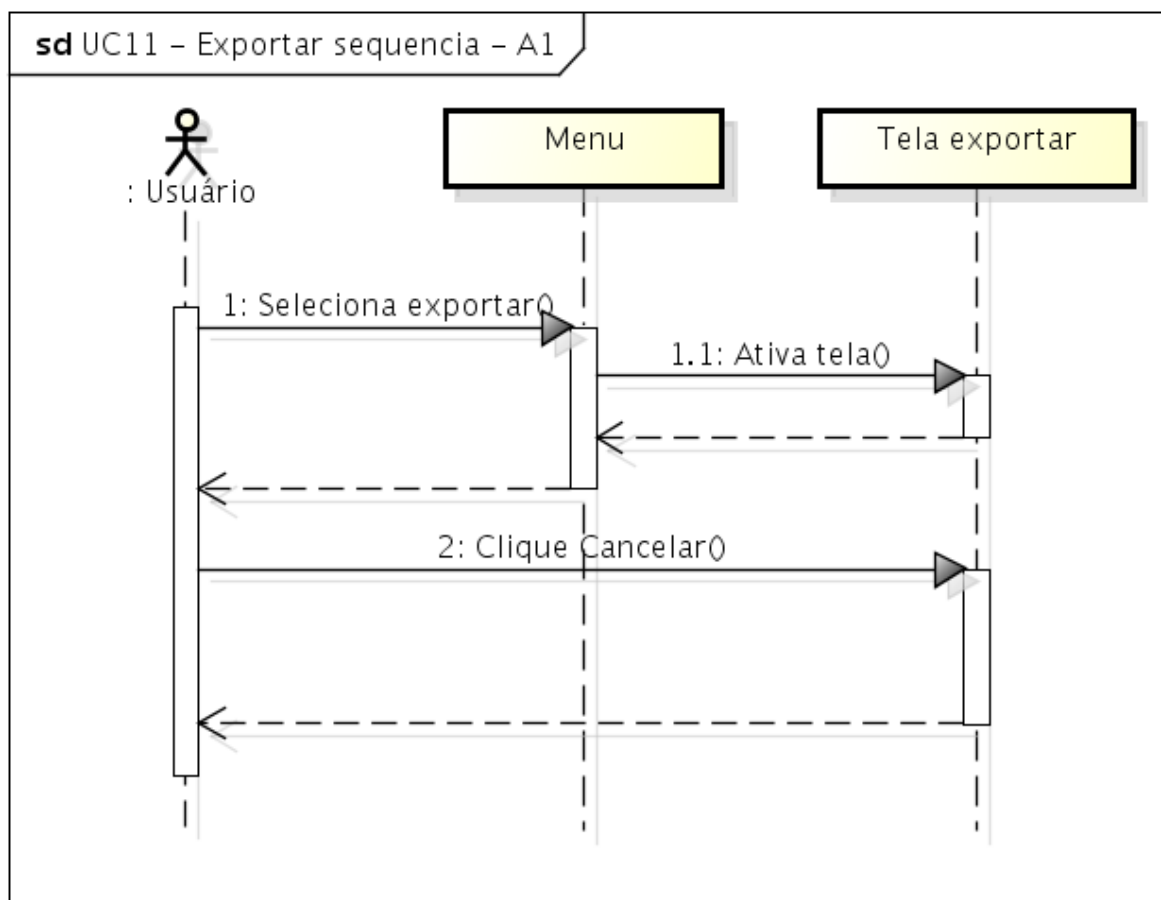
Figura 50 – Diagrama de Sequência – Visualizar alinhamento - Principal

## UC11 – Exportar sequência



powered by astah®

Figura 51 – Diagrama de Sequência – Exportar sequência – Principal



powered by astah\*

Figura 52 – Diagrama de Sequência – Exportar sequência – A1

## UC12 – Agrupar sequências

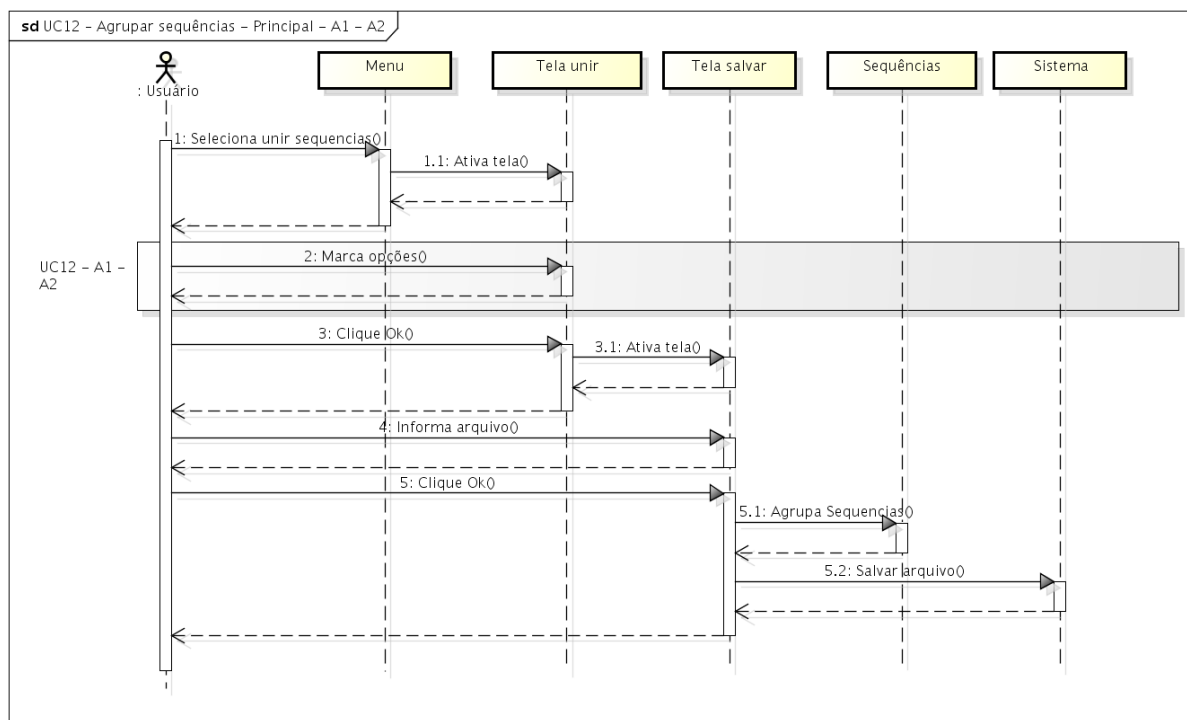


Figura 53 – Diagrama de Sequência – Agrupar sequências – Principal – A1 – A2

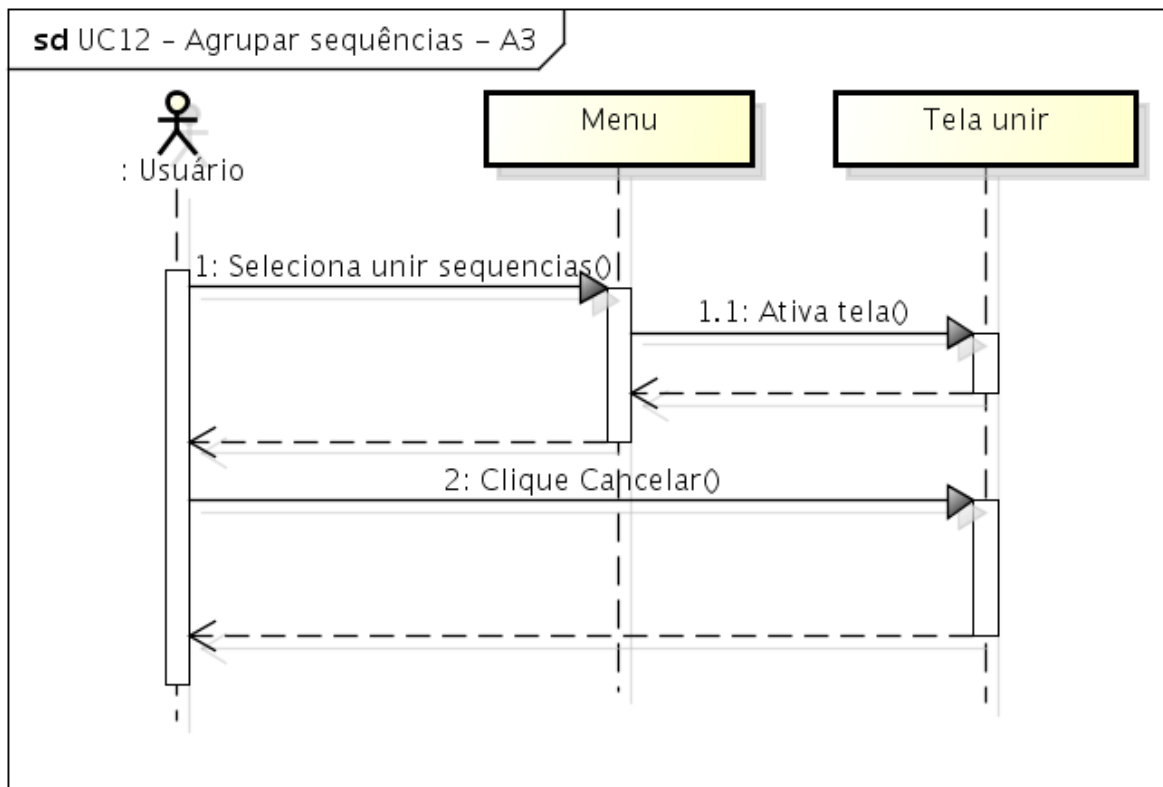
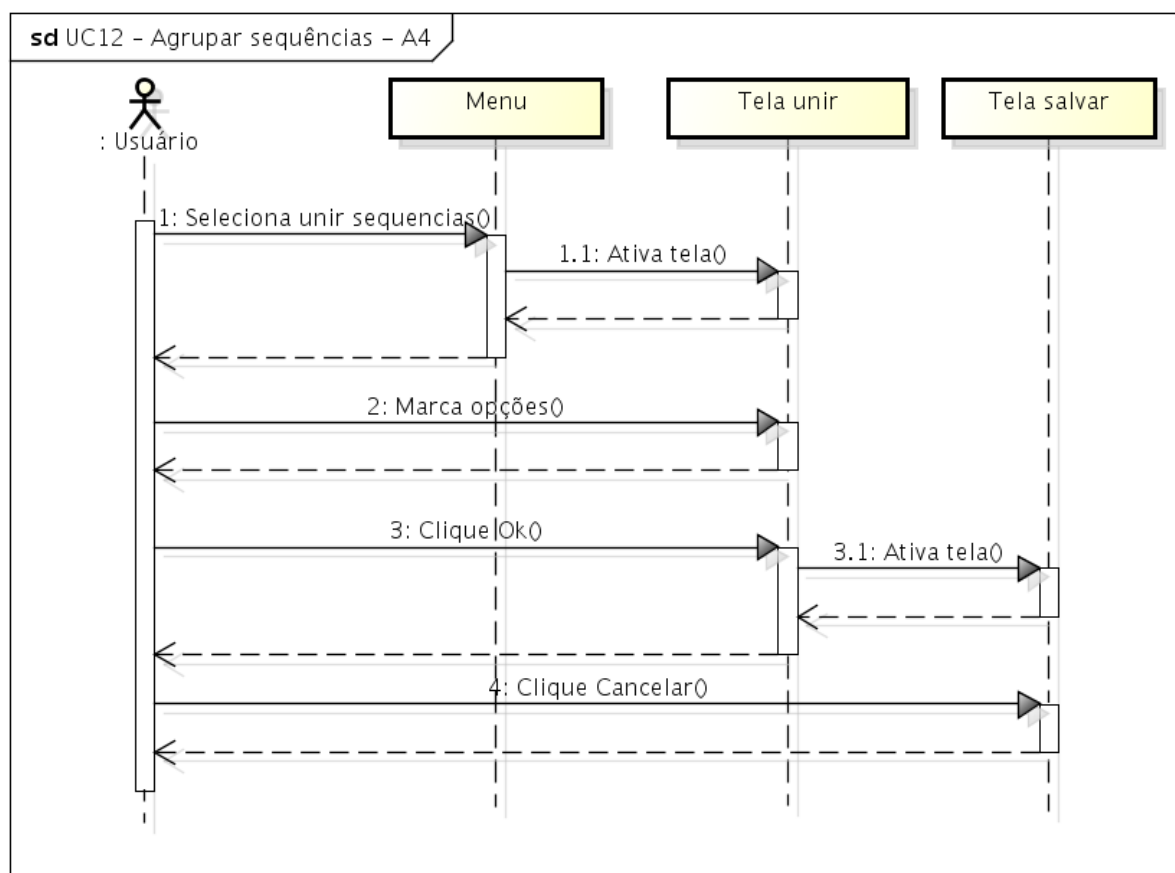


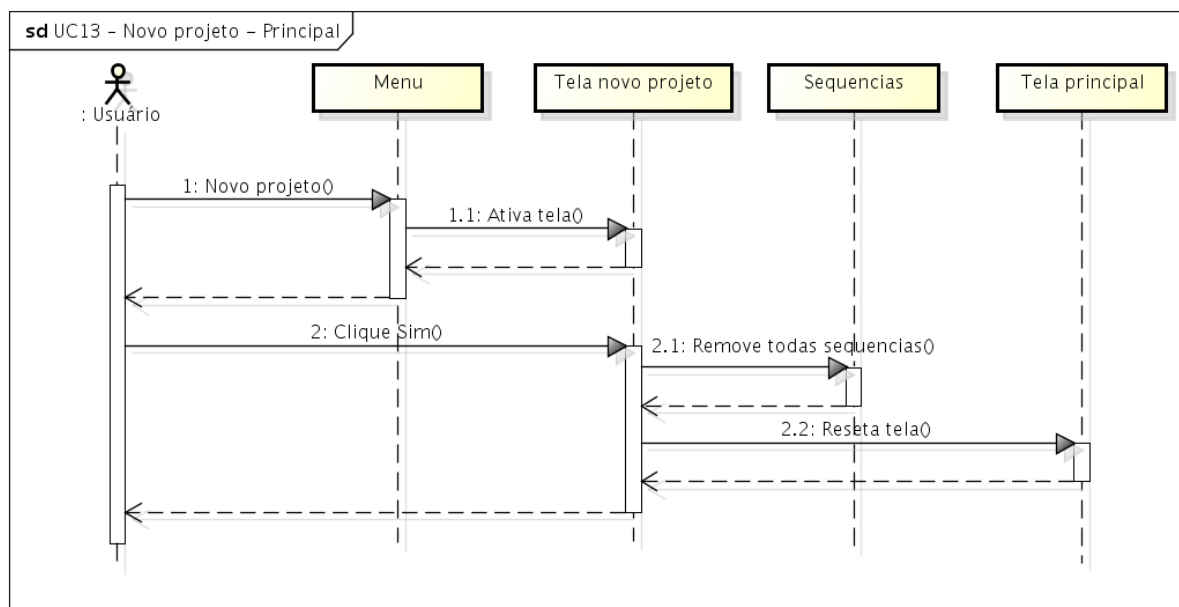
Figura 54 – Diagrama de Sequência – Agrupar sequências – A3



powered by astah®

Figura 55 – Diagrama de Sequência – Agrupar sequências – A4

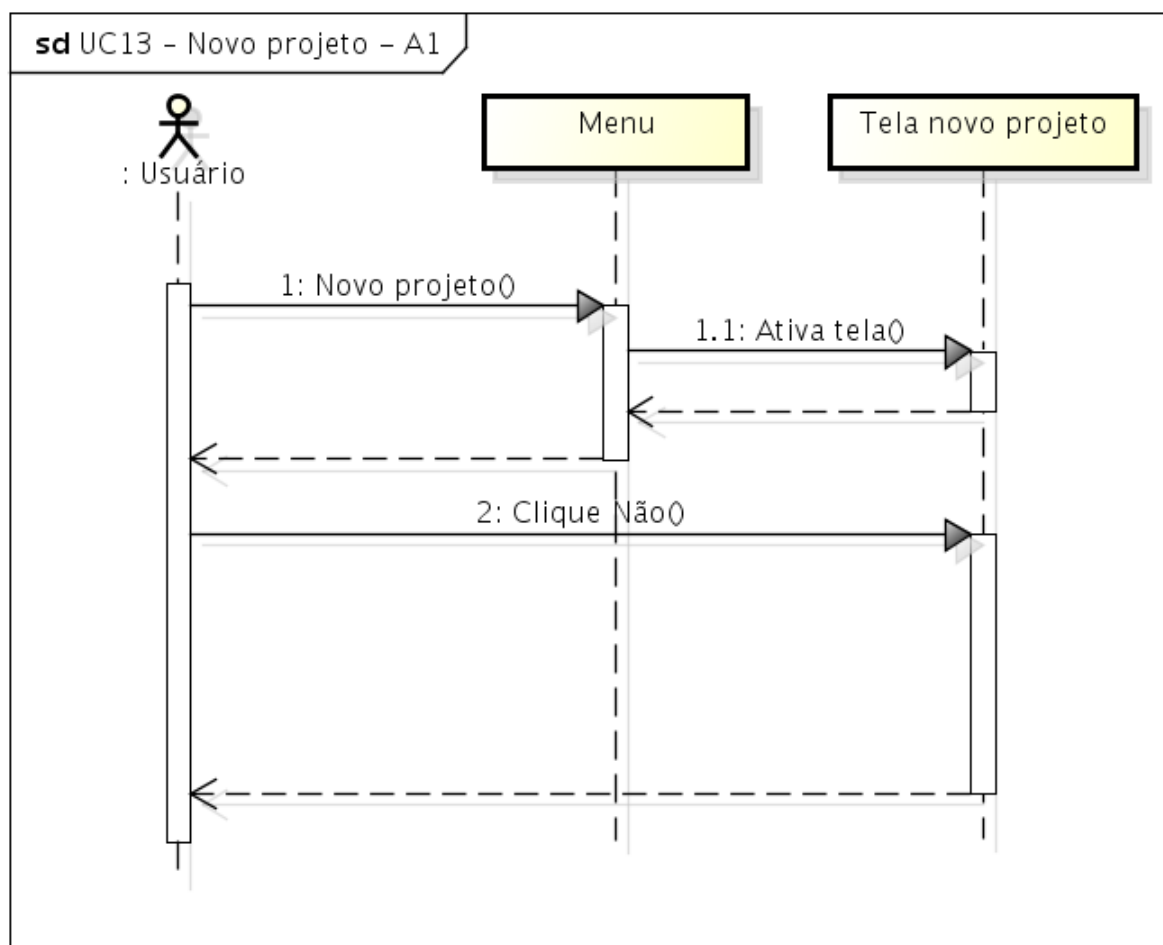
## UC13 – Novo projeto



powered by astah®

Figura 56 – Diagrama de Sequência – Novo projeto – Principal

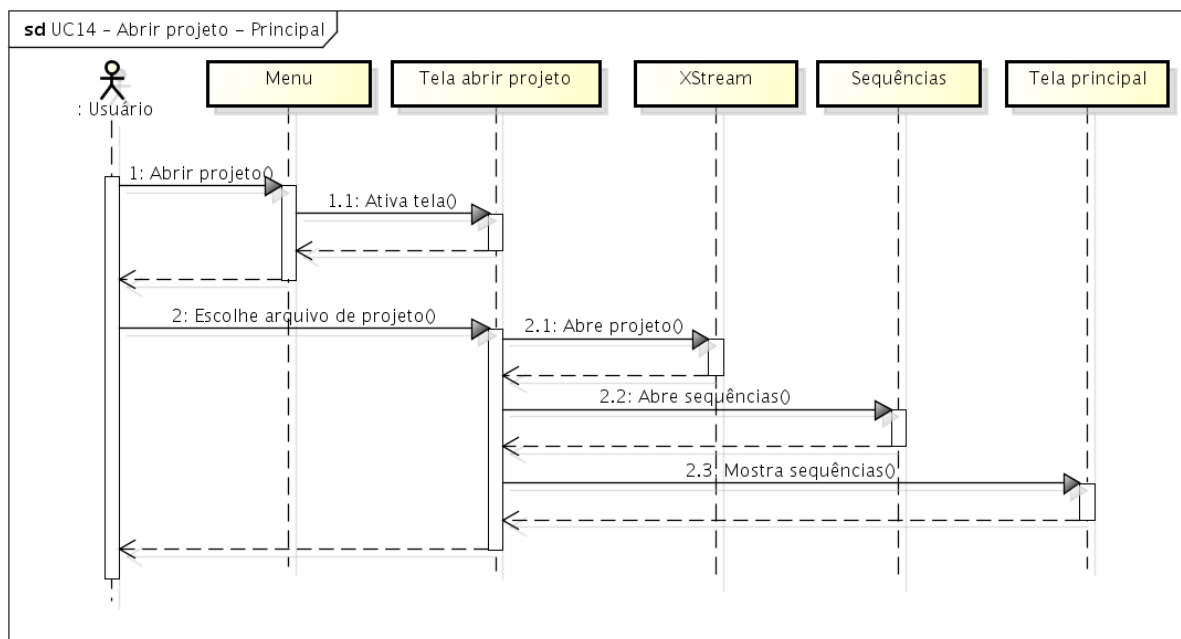




powered by astah\*

Figura 57 – Diagrama de Sequência – Novo projeto – A1

## UC14 – Abrir projeto



powered by astah®

Figura 58 – Diagrama de Sequência – Abrir projeto – Principal

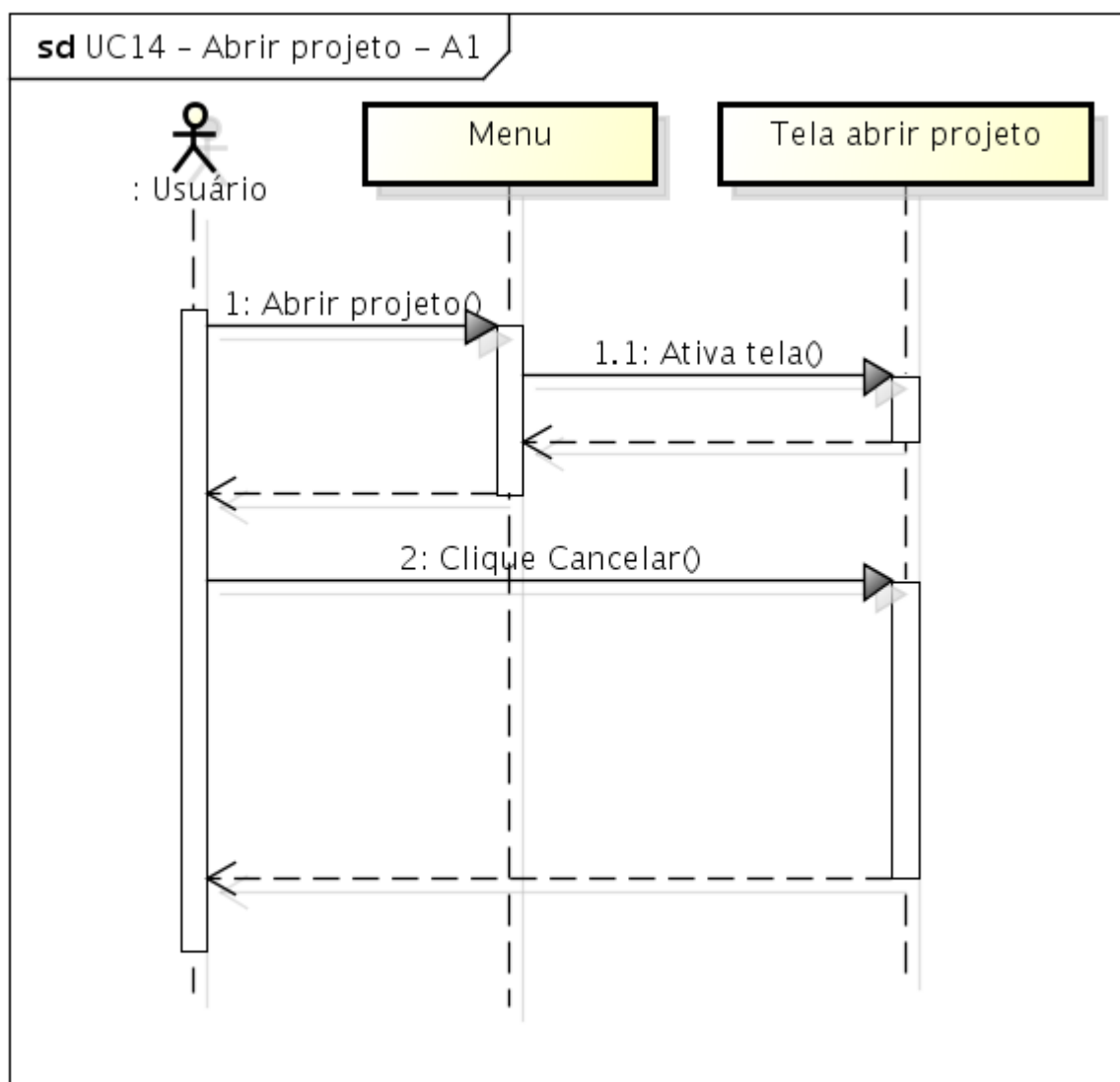
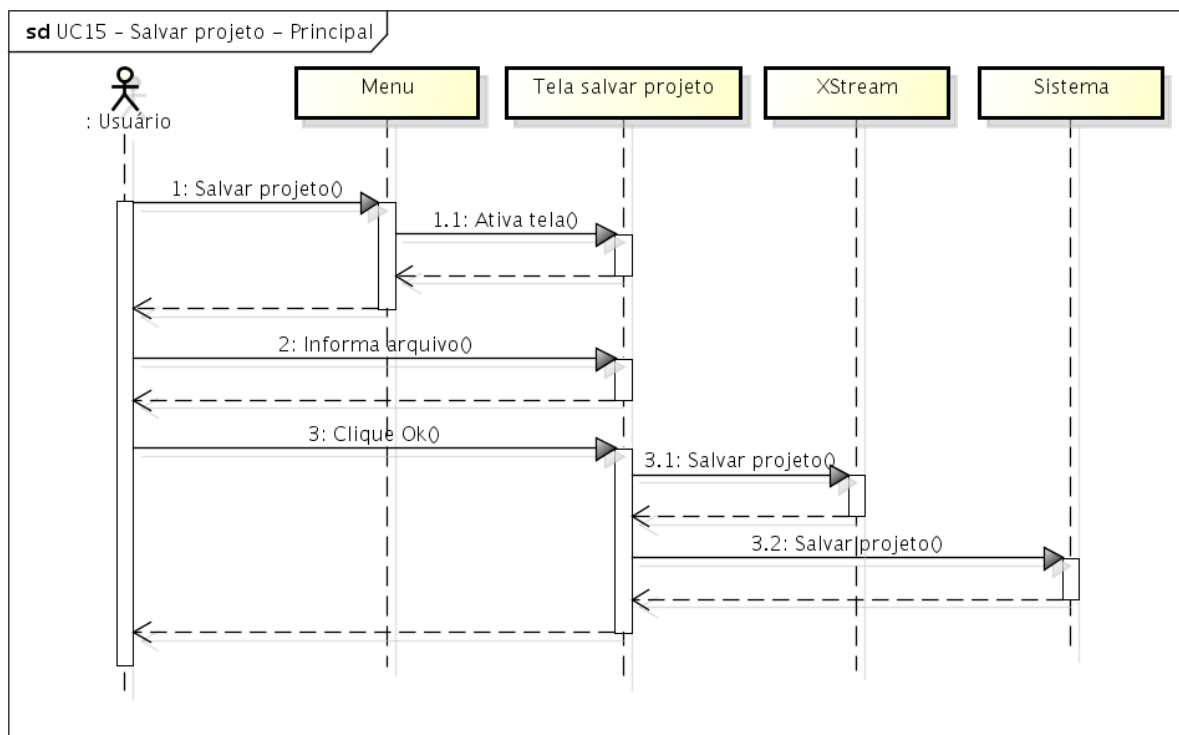
powered by astah<sup>®</sup>

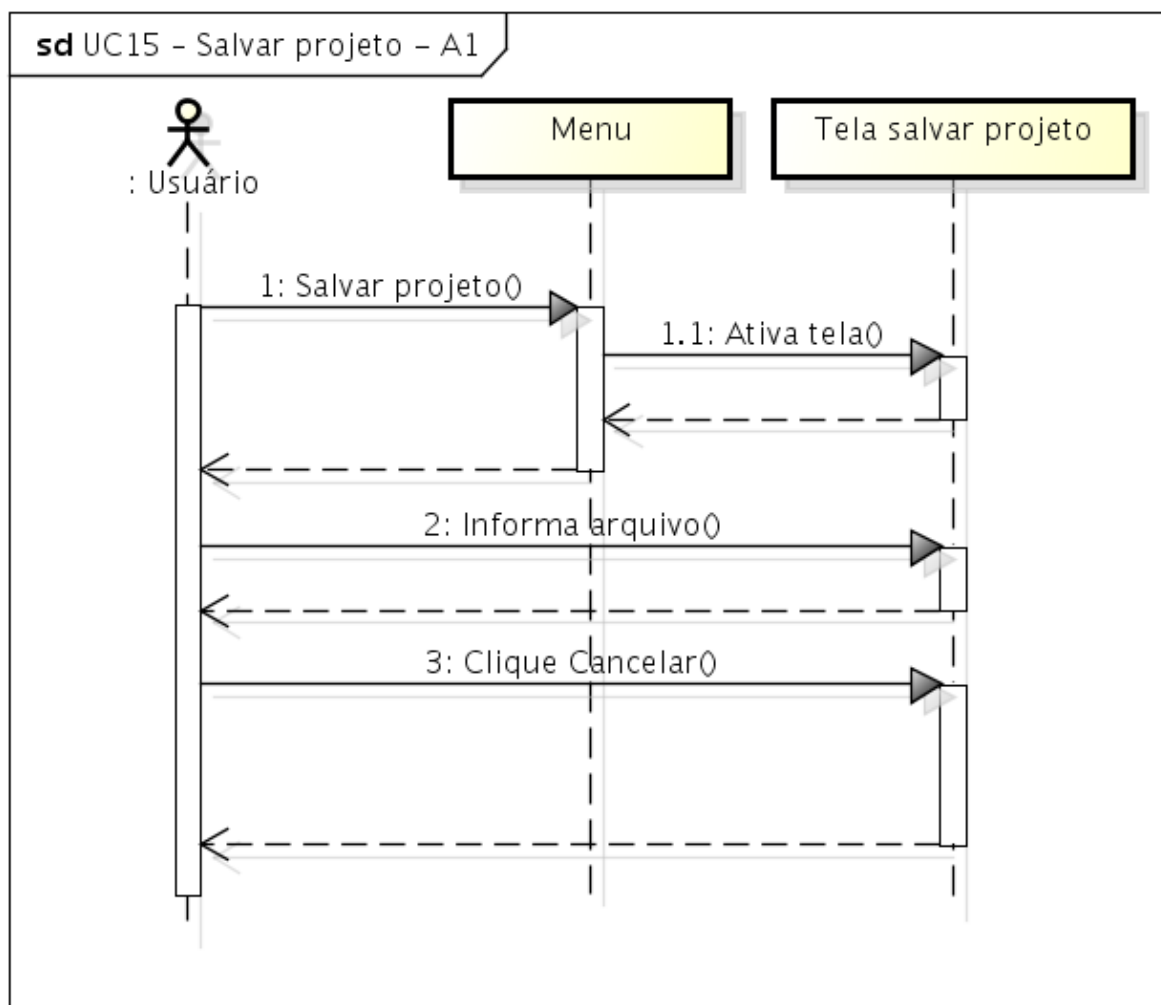
Figura 59 – Diagrama de Sequência – Abrir projeto – A1

## UC15 – Salvar projeto



powered by astah®

Figura 60 – Diagrama de Sequência – Salvar projeto - Principal



powered by astah\*

Figura 61 – Diagrama de Sequência – Salvar projeto – A1

## **5 PGP – PLANO DE GERENCIAMENTO DE PROJETO**

### **1 INTRODUÇÃO**

#### **1.1 ESCOPO E PROPÓSITO DO DOCUMENTO**

Este documento reúne os procedimentos necessários para o desenvolvimento de um software na área de bioinformática: JFinisher - finalizador de montagem e editor de sequências biológicas. Aqui serão descritos a forma de gerenciamento e organização necessários para a execução do projeto.

#### **1.2 OBJETIVOS DO PROJETO**

##### **1.2.1 OBJETIVOS**

O projeto do software JFinisher tem como objetivo gerar uma ferramenta capaz de auxiliar alguns processos realizados em laboratórios de bioinformática, na área de manipulação de sequências biológicas. Terá como objetivo principal unificar funcionalidades frequentemente utilizadas para facilitar e auxiliar na finalização de montagem de genomas recém-sequenciados.

##### **1.2.2 FUNÇÕES PRINCIPAIS**

A partir da entrada de arquivos texto padrão (Fasta ou GenBank), sendo um arquivo de referências e diversos arquivos de sequências, o software deverá:

- Fornecer uma interface navegável pelas sequências abertas, multilíngue (português e inglês, inicialmente)
  - Realizar alinhamentos entre sequências
  - Gerar sequência consenso
  - Editar e manipular sequências abertas
  - Salvar e abrir projetos em andamento
  - Exportar dados gerados

### 1.2.3 RESTRIÇÕES TÉCNICAS E ADMINISTRATIVAS

O projeto possui tempo de execução curto, um período de 5 meses. Por este motivo não é possível aplicar todas as funcionalidades necessárias para o problema de finalização de genomas recém-sequenciados. Esta limitação pode ser contornada com uma continuidade no projeto. Há uma limitação técnica no conhecimento da área de biologia molecular e bioinformática dos participantes, que terão que dividir o tempo de projeto e desenvolvimento com tempo de estudo dos conceitos e problemas da área.

## 2 ESTIMATIVAS DO PROJETO

### 2.1 DADOS HISTÓRICOS USADOS NAS ESTIMATIVAS

As estimativas do software foram baseadas em experiências profissionais dos envolvidos na área de desenvolvimento de software e planejamento. Para avaliar a complexidade de cada item, utilizamos valores padrão e nos remetemos aos trabalhos que foram previamente realizados.

A margem de erro desta estimativa é média/alta, pois os envolvidos não possuem experiência no desenvolvimento do softwares para bioinformática.

### 2.2 TÉCNICAS DE ESTIMATIVAS

A técnica utilizada para estimar o tempo para realização do software foi a análise de pontos por caso de uso. Esta técnica baseia as estimativas nos casos de uso e atores utilizados na modelagem do projeto (imagem UC).

Os elementos do sistema são classificados e atribuídos a pontuações conforme sua complexidade. Estes valores geram pontos de caso de uso não ajustados. Também são contabilizados fatores de complexidade técnica (características que podem atribuir dificuldades no desenvolvimento do software) e ambiental (nível de experiência dos profissionais envolvidos). Todos estes valores

calculados resultam em pontos de caso de uso ajustados. Utilizando a margem de 28 homens-hora, é estimada a duração do projeto.

## 2.3 ESTIMATIVAS

### Total de Pesos não Ajustados de Atores

<b>Atores</b>	<b>Nível de complexidade</b>
Usuário	3
BioInfoCommons	1
JcontigSort	1
Jaligner	1
<b>Total</b>	
Simples (Total: 3 – Peso: 1)	3
Médio (Total: 0 – Peso: 2)	0
Complexo (Total: 1 – Peso: 3)	3
<b>TPNAA</b>	<b>6</b>

### Total de Pesos não Ajustados dos Casos de Uso

<b>Caso de uso</b>	<b>Nível de complexidade</b>
Abrir Arquivo	1
Abrir arquivo referência	1
Abrir/Adicionar arquivo(s) de sequência/contigs	2
Remover sequência	2
Exportar sequência	2
Agrupar sequências	3
Navegar sequências	2
Localizar/lr para posição	1
Alinhar sequência	3
Visualizar alinhamento	3
Editar sequência	2
Excluir	2
Inserir	2
Novo projeto	1
Abrir projeto	3
Salvar projeto	3
<b>Total</b>	
Simples (Total: 4 – Peso: 1)	4
Médio (Total: 7 – Peso: 2)	14
Complexo (Total: 5 – Peso: 3)	15
<b>TPNAUC</b>	<b>33</b>

### Fator de complexidade técnica

<b>Descrição</b>	<b>Peso</b>	<b>Valor</b>	<b>Total</b>
Sistemas Distribuídos	2	2	4
Desempenho da Aplicação	1	5	5
Eficiência do usuário final	1	4	4



Processamento interno complexo	1	5	5
Reusabilidade do código	1	2	2
Facilidade de Instalação	0,5	3	1,5
Usabilidade	0,5	5	2,5
Portabilidade	2	1	2
Manutenibilidade	1	1	1
Concorrência	1	0	0
Características especiais de segurança	1	4	4
Acesso direto para terceiros	1	0	0
Facilidades especiais de treinamento	1	2	2
<b>Soma</b>			33
<b>FCT = <math>0,6 + (0,01 * Soma) = 0,93</math></b>			

#### Fator de complexidade ambiental

<b>Descrição</b>	<b>Peso</b>	<b>Valor</b>	<b>Total</b>
F1 Familiaridade com o processo de desenvolvimento de software	1,5	2	3
F2 Experiência na aplicação	0,5	2	1
F3 Experiência com OO, na linguagem e na técnica de desenvolvimento	1	3	3
F4 Capacidade do líder de análise	0,5	3	1,5
F5 Motivação	1	5	5
F6 Requisitos estáveis	2	0	0
F7 Trabalhadores com dedicação parcial	-1	4	-4
F8 Dificuldade na linguagem de programação	-1	2	-2
<b>Soma</b>			7,5
<b>FCA = <math>1,4 + (-0,03 * Soma) = 1,175</math></b>			

Pontos por caso de uso não ajustados:

- $PCUNA = TPNA A + TPNA UC = 39$

Pontos por caso de uso ajustados:

- $PCUA = PCUNA * FCT * FCA = 42,61725$

Estimativa de horas:

- $28hh * 42,61725 = 1193,283horas = 132,587 dias$

### 3 RISCOS DO PROJETO

#### 3.1 ANÁLISE DOS RISCOS

##### 3.1.1 IDENTIFICAÇÃO

1. Identificação de bibliotecas: o projeto irá necessitar de bibliotecas para o desenvolvimento de algumas funcionalidades. São bibliotecas que implementam algoritmos de alinhamento entre sequências e bibliotecas que manipulem arquivos da área.
2. Pessoal: a equipe possui apenas 2 integrantes e é dependente das suas funções para o projeto inteiro, tendo um cronograma justo para a realização das tarefas.
3. Linguagem de programação: o conhecimento avançado em alguns assuntos da linguagem utilizada (Java) será necessário, sendo necessário estudo prévio. É possível também que a linguagem não consiga suprir uma necessidade da maneira como esperado.
4. Tempo: o tempo destinado ao planejamento e desenvolvimento do projeto não pode ser ultrapassado.
5. Conhecimento limitado na área: por ser um software de bioinformática, é necessário um conhecimento de alguns conceitos previamente para execução do projeto.
6. Desempenho: o sistema irá demandar intenso uso do hardware e deve ser planejado para conseguir atingir esta meta. Isto pode gerar atrasos na análise de complexidade dos algoritmos utilizados para que o software seja executado sem problemas.

##### 3.1.2 AVALIAÇÃO DOS RISCOS

1. Identificação de bibliotecas:

Data limite: 28/10/2011

Consequência: Atraso no início do desenvolvimento de algumas funcionalidades.

Ação: Encontrar funcionalidades alternativas ou programar a biblioteca.

Monitoramento: ao início do desenvolvimento e na implementação das bibliotecas

Probabilidade: Muito baixa

Impacto: Muito alto

## 2. Pessoal:

Data limite: Todo o projeto

Consequência: Possível atraso em algum estágio

Ação: Alterar cronograma conforme dificuldades inesperadas acontecem, possibilitando que os estágios sejam finalizados.

Monitoramento: Constante, semanal

Probabilidade: Moderada

Impacto: Baixo / Moderado

## 3. Linguagem de programação:

Data limite: Todo o projeto

Consequência: Possível atraso em algum estágio, impossibilidade de realizar funcionalidades

Ação: rever a funcionalidade de acordos com as limitações da linguagem, buscar métodos alternativos. Manter estudo na área da linguagem de programação utilizada para evitar outros riscos

Monitoramento: um estágio de antecedência do desenvolvimento

Probabilidade: Moderada

Impacto: Alto

## 4. Tempo:

Data limite: 20/02/2011

Consequência: Atraso na entrega

Ação: Simplificar estágio e/ou funcionalidade para ajustar tempo de desenvolvimento. Dedicar mais tempo ao projeto

Monitoramento: Constante verificação do cronograma

Probabilidade: Moderada

Impacto: Alto

#### 5. Conhecimento limitado na área:

Data limite: Todo o projeto

Consequência: Não compreensão dos problemas, dificuldades no desenvolvimento

Ação: Realizar estudo prévio, identificar todo o conteúdo que envolva o projeto, reunião com orientador para dúvidas

Monitoramento: Constante

Probabilidade: Alta

Impacto: Baixo

#### 6. Desempenho:

Data limite: Ao final de cada estágio

Consequência: remoção de funcionalidade, software com tarefas lentas, atraso na entrega

Ação: Identificar os algoritmos críticos e revisar bibliografia para melhorar o desempenho. Utilizar outros métodos

Monitoramento: Durante as tarefas de desenvolvimento do projeto

Probabilidade: Alta

Impacto: Alto no desempenho, moderado no atraso

### 3.2 ADMINISTRAÇÃO DOS RISCOS

#### 3.2.1 OPÇÕES PARA EVITAR OS RISCOS

1. Identificação de bibliotecas: pesquisa com antecedência de bibliotecas que realizam as funções desejadas. Encontrar várias opções, tendo redundância e opção de escolha. Verificar bibliotecas comuns na área, as mais utilizadas e recentes, com documentação atualizada.
2. Pessoal: planejar cronograma com mais tempo para tarefas críticas.
3. Linguagem de programação: revisar bibliografia da linguagem nos assuntos que serão mais utilizados. Analisar programas semelhantes e suas técnicas.
4. Tempo: ser cauteloso na distribuição de tempo. Integrantes devem estar cientes que podem precisar de mais tempo do que o normal para finalização do projeto, abrindo mão de outras atividades.

5. Conhecimento limitado na área: realizar estudo e ter visão geral do problema a ser desenvolvido. Procurar assistência com orientador do projeto e profissionais da área para sanar dúvidas. Dividir estudos por estágios, conforme necessidades aparecem.
6. Desempenho: verificar técnicas de softwares semelhantes. Procurar métodos alternativos para solução de problemas. Rever funcionalidade.

### 3.2.2 PROCEDIMENTOS DE MONITORAÇÃO DOS RISCOS

Serão realizadas reuniões semanais para avaliar o andamento e avaliar possíveis riscos no decorrer do projeto. Desta maneira será mais fácil a rápida identificação do problema, encontrando uma possível solução, que não resulte em atras nem problemas no software.

4 CRONOGRAMA

4.1 WORK BREAKDOWN STRUCTURE– DIVISÃO DE TRABALHO NO PROJETO

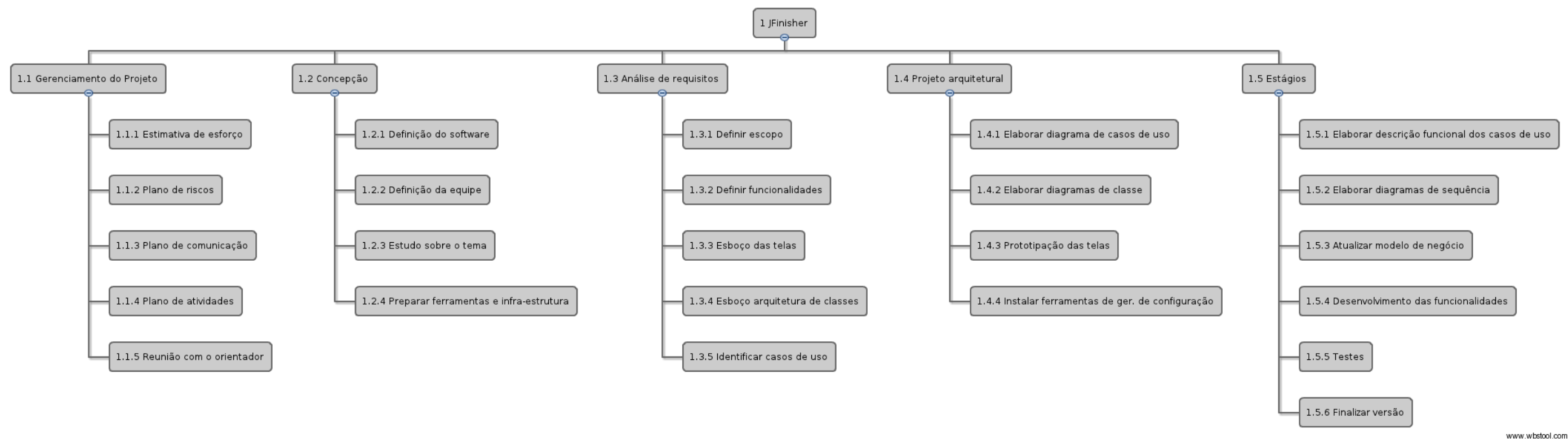


Figura 62 - WORK BREAKDOWN STRUCTURE

4.2 GRÁFICO DE GANTT

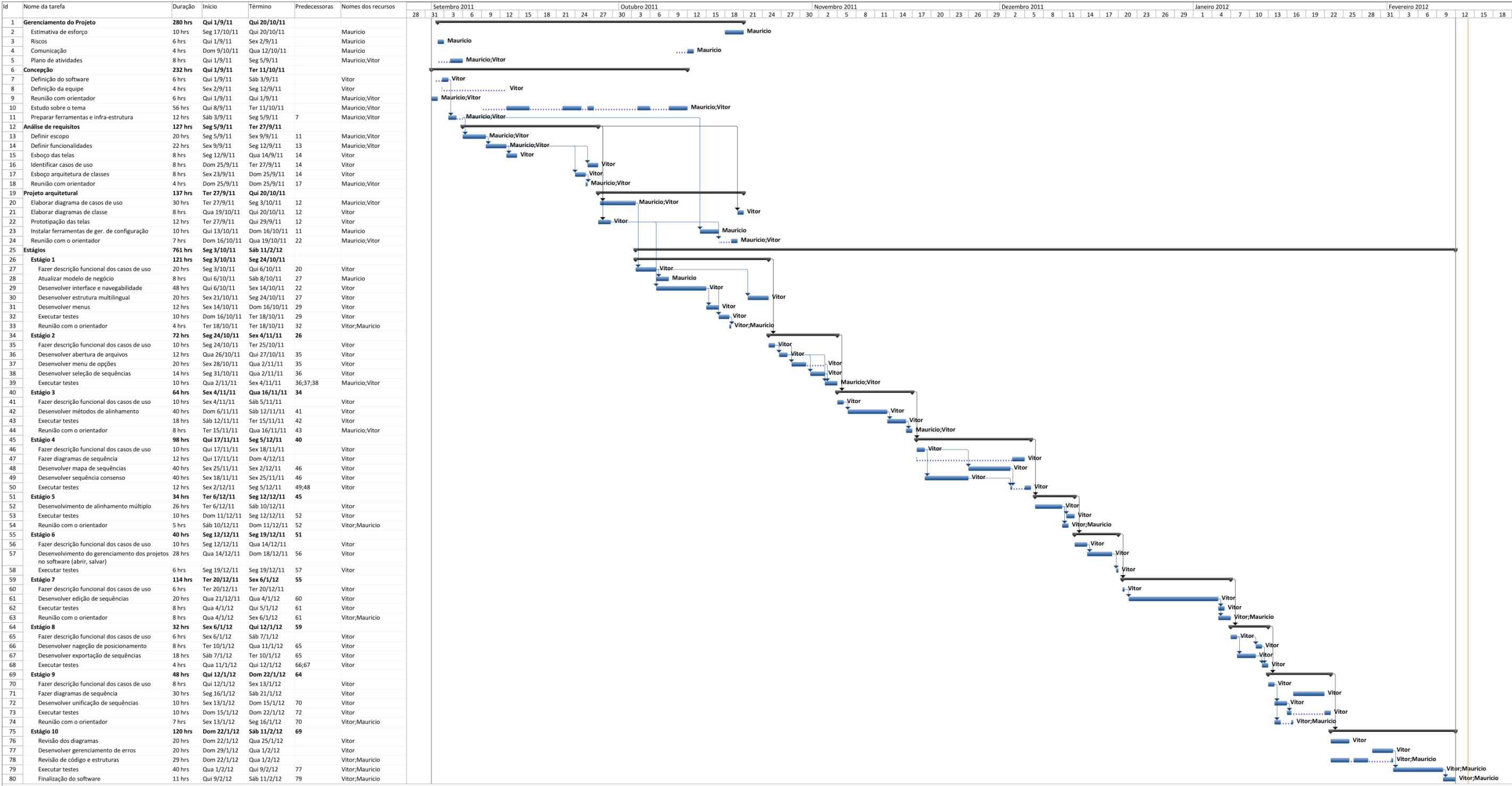


Figura 63 - GANTT

## **5 RECURSOS DO PROJETO**

### **5.1 PESSOAL**

A equipe será formada por 2 membros: Vitor Cedran Piro e Mauricio Tarabella Galvão, tendo o Professor Dieval Guizelini como professor orientador.

### **5.2 HARDWARE E SOFTWARE**

Para realização do projeto, desenvolvimento, testes e documentação serão utilizados 2 notebooks:

- Processador Core 2 Duo (1.8GHz), Memória RAM 3GB, HD 250GB
- Processador Quad Core, Memória RAM 3GB, HD 1TB

O programa será desenvolvido na plataforma JAVA, JDK versão 6, utilizando os seguintes softwares (para desenvolvimento, planejamento e documentação):

- Sistema operacional: Linux Mint 11.04 / Ubuntu 10.04
- IDE de desenvolvimento e testes: NetBeans versão 7.0.1
- Documentação: Microsoft Office Professional Plus 2010\*, Google Docs
- Diagramas: Astah Community 6.5.1, Microsoft Project Professional 2010\*, WBSTool.com, Planner v0.14.6

\* Obs: Os softwares que necessitam de licença paga foram utilizados em suas respectivas versões de avaliação, com licença para estudantes.

## **6 ORGANIZAÇÃO DO PESSOAL**

### **6.1 ESTRUTURA DA EQUIPE**

Vitor será o analista e desenvolvedor do sistema, além de contribuir na parte estrutural e de testes. Maurício se encarrega da gerencia e acompanhamento do



projeto e construção de diagramas. Dieval como professor orientador, irá auxiliar no andamento do projeto e possíveis dúvidas.

## 7 PLANO DE COMUNICAÇÃO

Grupos interessados	Foco	O que este grupo precisa saber	Método	Quando	Observação
Equipe	Compartilhar Conhecimento  Dificuldades e aprendizados	Quais as dúvidas?  Onde pesquisar?  O que pesquisar?	Troca de conteúdo estudado	Diariamente	Quem tiver algo a compartilhar (dúvidas, descobertas, realizações) compartilha via e-mail ou outro meio disponível
Equipe	Feedback das atividades	O que desenvolveu	Reunião de 2 horas	Semanalmente	Envolvidos enviam e-mails confirmando presença.
Equipe + Professor orientador	Tirar dúvidas  Verificar pendências  Verificar funcionalidades	Dificuldades e problemas encontrados  Soluções encontradas	Reunião de 2 horas	Quinzenalmente	Responsável agenda horário com professor orientador
Equipe + Usuário especializado	Testar a aplicação com dados reais de laboratório	Usar o software  Obter os arquivos para teste	Encontro em laboratório da UFPR para uso do software	Dependendo da disponibilidade do usuário, reunião ocorre na fase final do projeto	Responsável agenda horário com usuário especializado

## 8 MECANISMOS DE RASTREAMENTO E CONTROLE

Os controles do projeto são de responsabilidade de todos os envolvidos. Os aspectos a serem controlados estão indicados abaixo.

## 8.1 VERIFICAÇÃO E CONTROLE DO ESCOPO

Para a verificação do escopo serão utilizados check lists em reuniões periódicas a fim de manter o projeto nas conformidades do escopo. Possíveis alterações de projeto que possam impactar no escopo deverão ser atualizadas no check-list para que a documentação seja refatorada.

## 8.2 CRONOGRAMA

O controle do cronograma será executado com o auxílio da lista de atividades e o do gráfico de Gantt. Com isto será possível o acompanhamento das atividades bem como seus possíveis atrasos, permitindo o remanejamento rápido de recursos para evitar atrasos excessivos. Como o projeto se trata de um trabalho acadêmico, existe um limite nos recursos humanos empregados, não existindo a possibilidade de agregação de novos recursos quando necessário, isso torna o cronograma um risco.

## 8.3 QUALIDADE

A verificação da qualidade do projeto será obtida por testes periódicos do software pelo professor orientador e por uma equipe especialista por ele indicada. Assim teremos retorno da qualidade e andamento do desenvolvimento.

## 8.4 GERENCIAMENTO DA EQUIPE DO PROJETO

Será realizado por cada membro da equipe uma vez que não existe uma hierarquia dentre os participantes do projeto. Os membros deverão observar uns aos outros para garantir o comprometimento e o cumprimento das exigências do projeto. A saída repentina de membros deverá ser sanada com a agregação de membros substitutos o mais rápido possível para que não aconteçam grandes impactos no cronograma.

## 8.5 RELATÓRIO DE DESEMPENHO

O desempenho do andamento do projeto será analisado em reuniões periódicas. Assim será possível analisar as metas estabelecidas e atingidas para cada um dos indicadores do projeto.

## 8.6 GERENCIAMENTO DAS PARTES INTERESSADAS

Reuniões deverão acontecer com o orientador do projeto para que se garanta o sucesso na entrega do mesmo. As opiniões e solicitações do coordenador serão de grande relevância uma vez que o projeto será avaliado por uma banca de professores ao final do prazo estimado.

## 8.7 RISCOS

O propósito da monitoração dos riscos é determinar se as respostas aos riscos estão sendo implementadas como planejadas e tão efetivas quanto esperadas. Também visa se devem ser desenvolvidas novas respostas, bem como se as premissas do projeto ainda são válidas, e se a exposição ao risco mudou frente ao seu estado anterior. O acompanhamento deve observar se um gatilho de risco ocorreu e se as políticas e procedimentos adequados estão sendo seguidos. O relatório de riscos será avaliado periodicamente nas reuniões para este fim e as ações serão conduzidas de acordo com as conclusões e decisões das reuniões de acompanhamento.

## 8.8 MUDANÇA DE ESCOPO

Qualquer alteração que se fizer necessária e que implique na alteração de um dos documentos do projeto, será considerada como mudança de escopo. A solicitação deverá ser analisada por todos os membros da equipe para medir o impacto da mudança no projeto e definição da recomendação a ser seguida. A implementação de uma mudança somente acontecerá, se recomendada pela equipe e aprovada pelo orientador do projeto.