

Leticia do Nascimento

**LAPOGE DATABASE - ESTRUTURA RELACIONAL DE BANCO
DE DADOS BIOLÓGICOS PARA ARMAZENAMENTO E
SISTEMATIZAÇÃO DE INFORMAÇÕES GENÉTICAS E
EPIDEMIOLÓGICAS**

Trabalho de Conclusão de Curso
apresentado como requisito para
cumprimento da disciplina TCC II
(BIO7016) do currículo do Curso de
Graduação em Ciências Biológicas da
Universidade Federal de Santa Catarina.

Orientadora: Prof^a Dr^a Andrea Rita
Marrero

Florianópolis

2016

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

do Nascimento, Leticia

LAPOGE Database - Estrutura relacional de banco de dados biológicos para armazenamento e sistematização de informações genéticas e epidemiológicas / Leticia do Nascimento ; orientadora, Andrea Rita Marrero - Florianópolis, SC, 2016. 88 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas. Graduação em Ciências Biológicas.

Inclui referências

1. Ciências Biológicas. 2. Genética Epidemiológica. 3. Organização de dados. 4. Marcadores genéticos . I. Marrero, Andrea Rita. II. Universidade Federal de Santa Catarina. Graduação em Ciências Biológicas. III. Título.

Leticia do Nascimento

**LAPOGE *DATABASE* - ESTRUTURA RELACIONAL DE BANCO
DE DADOS BIOLÓGICOS PARA ARMAZENAMENTO E
SISTEMATIZAÇÃO DE INFORMAÇÕES GENÉTICAS E
EPIDEMIOLÓGICAS**

Este Trabalho de Conclusão de Curso foi julgado adequado para o cumprimento da disciplina TCCII (BIO7016) e aprovado em sua forma final pelo Banca Examinadora.

Florianópolis, 8 de Julho de 2016

Prof^a Dr^a Maria Risoleta F. Marques
Coordenadora do Curso de Ciências Biológicas

Banca Examinadora:

**Prof^a Dr^a Andrea
Rita Marrero
Orientadora**

Prof^a Dr^a Ilíada
Rainha de Souza

Prof^a Dr^a Sara
Emelie Löfgren

Para José e Sandra, pais e eternos
guardiões;

E aos meus preciosos amigos.

AGRADECIMENTOS

Não poderia deixar de começar agradecendo aos meus pais, pelas diversas privações as quais passaram para me garantir uma vida de qualidade e com uma ótima educação. Pelo incentivo constante para que eu tivesse algo que nunca puderam ter e pelo amor e carinho que se faz tão difícil demonstrar em nossa família, mas que sem dúvidas, não nos falta.

Aos meus colegas de curso, e em especial: Leonardo e as diversas caronas e conversas reclamando da vida; a Tainha e Corvo, pela amizade que tivemos durante nossos primeiros semestres, espero e torço sinceramente por sua felicidade; a Andressa, e minha única amizade feminina duradoura na faculdade.

A Lucas, um cara complicado que vive sob uma pele de lobo. Por me proporcionar boa parte das melhores e mais docemente dolorosas memórias que possuo. Aos inúmeros livros e filmes que compartilhamos, a madrugada em que eu não aguentava mais e que você veio zelar por mim enquanto digitava um trabalho chorando por puro desespero e aos doces e besteiras que você me trouxe quando as coisas estavam ruins. E quando estavam boas também.

Rodrigo, o irmão que eu pude escolher ao decorrer da vida. Obrigada por acompanhar e por me apontar sempre que possível os meus defeitos e os meus crescimentos, por sempre me incentivar a ser o meu melhor e por sempre me lembrar que bom mesmo é ser bom.

Diogo, o senhor do tempo e dos mundos. Por me salvar da minha própria vida me dando a oportunidade de ser outras pessoas em mundos diferentes e passando por provações diferentes. Há quem diga que RPG é coisa de criança, coitados, apenas não sabem a provação que ter um mestre e um amigo como você pode ser na vida de alguém, ou em todas as vidas que você escolhe ter.

Fábio, a parte de mim que na viagem pra vida se perdeu e foi parar em outro alguém. Pela amizade incondicional e independente de distancias e barreiras. Por me fazer sentir que eu tinha uma outra família em outro estado, por me receber de braços abertos após um congresso e por me apoiar sempre.

Axel e a cara de maluco que tantas vezes me fez rir quando eu mais precisava. Por todas as palavras de apoio, pelo esforço em me fazer reaprender a gostar de mim e por se mostrar sempre disponível pra me estender a mão. Pelo notebook que salvou boa parte das tabelas de

resultados do meu TCC e por dificilmente me deixar almoçar sozinha durante esse semestre excruciante, sua presença diária fez toda diferença.

A Dr. Andrea Rita Marrero, que aceitou a difícil tarefa de me orientar. Pela confiança, depositando em minhas mãos um trabalho que há tantos anos desejava pessoalmente implementar e por todos os “TATU’s” quando eu estava parecendo o urso do pica pau, Obrigada! Por me ensinar muito mais do que evolução e genética, por ensinar na base do exemplo valores raros de encontrar hoje dia, és uma pessoa incrível e só posso agradecer por ter tido a oportunidade de estar por perto. Que a força esteja sempre com você!

Ao LAPOGE e todos que já fizeram parte do mesmo, por me receber de braços abertos e por emprestar sua história e trajetória, afim de produzir esse trabalho.

Gostaria de agradecer também a banca Dr^a Sara Emelie Löfgren e Dr^a Ilfada Rainha de Souza, por cordialmente aceitar participar deste trabalho.

À Universidade Federal de Santa Catarina. A todos os professores que passaram pela minha jornada, que me ensinaram muito sobre Biologia e ousou dizer, que ainda mais sobre a vida. A todos os servidores e técnicos administrativos e os demais que direta ou indiretamente fizeram parte de minha formação.

E a todas outras pessoas que passaram de alguma forma pela minha trajetória e que hoje já não se fazem mais presentes.

“Faça. Ou não faça. Não existe tentar”

(Yoda, Guerra nas Estrelas: O Império Contra-
Ataca, 1980)

RESUMO

O Laboratório de Polimorfismos Genéticos – LAPOGE da UFSC foi implantado há 22 anos e é base para a formação de alunos de graduação e pós-graduação, além de estágios e vivências. O impressionante volume de dados gerados neste tempo fez crescer a necessidade de um sistema de organização dos dados coletados para todos os participantes das diversas pesquisas executadas neste tempo. Neste sentido, este trabalho propõe a implementação de um banco de dados relacional adequado ao LAPOGE, além do levantamento dos trabalhos realizados e questionários utilizados para levantamento de informações do participante doador da amostra. Durante o levantamento de dados, foi possível encontrar um total de 49 produções. Dentre estas, 34 trabalhos de conclusão de curso de graduação, 12 dissertações de mestrado e 3 teses de doutorado, que utilizaram um total de 158 marcadores. Além disso, notou-se uma grande quantidade de questionários diferentes gerados ao longo dos trabalhos; para garantir uma maior precisão na coleta de dados e diminuir a margem de erro, foram levantadas todas as perguntas comuns utilizadas durante a coleta dos dados biológicos e criado um novo questionário unificado e padronizado. O LAPOGEDb (LAPOGE Database) foi tabulado para armazenar dados em MySQL com interface PHP. Os dados são criptografados e o acesso é restrito a usuários com cadastro autorizado. Está depositado no servidor da Universidade Federal de Santa Catarina que executa três cópias diárias de segurança (*backup*). A estrutura relacional foi organizada em dez abas e a interface final é do tipo “amigável” onde não é exigido ao usuário nenhum conhecimento de programação. Foram implementadas as abas de Dados Pessoais, Dados Familiares, Dados Biológicos, Dados Epidemiológicos, Dados médicos, Marcadores genéticos, Amostra biológica e as abas específicas de Artrite, Psoríase, Lúpus e Câncer de Mama. O sistema de cadastro permite o acesso às informações, mas mantém sigilosas aquelas previstas pelo Comitê de Ética. A página de busca (que será implementada) permitirá coletar informações combinadas de todas as categorias e emitirá *output* em formato de uso para programas de análises. Dado o volume de informações, o LAPOGEDb inicia uma nova forma de organização e consulta de dados produzidos pelo LAPOGE UFSC.

Palavras-chave: Genética Epidemiológica; organização de dados; marcadores genéticos

ABSTRACT

The LAPOGE – Genetic Polymorphism Laboratory – was created 22 years ago and is still a place for development of both undergraduated and graduated students. The great amount of data generated in this period generated the necessity to organize the data gathered by several research lines into a relational data bank suited to LAPOGE needs, where all questionnaires applied, information on sample donors and results of previous works carried on with those samples. During data compilation, 49 studies were retrieved. Among these, 34 were undergraduate monographies, 12 master dissertations and 3 doctoral thesis and a total of 158 genetic markers were analyzed. Several questionnaire layouts were used for each study, turning data padronization difficult and increasing the chance of sampling error. A new questionnaire was elaborated with standard questions to ensure a broad data sampling of donor information. The LAPOGED*b* (LAPOGE Database) was designed in MySQL with a PHP interface. Data is encrypted and access is restricted to authorized users. The Database is stored at the UFSC server, which executes three backups on a daily basis. The relational structure was organized in ten frames, and the interface is user-friendly, which means the user does not need any knowledge in programming. The frames include Personal data, Family data, biological data, epidemiological data, medical history, genetic markers, biological sample status and frames for the disease studied, as Arthritis, Psoriasis, SLE and Breast cancer. The registration system allows access to information but keeps private those required by the Ethics Committee and the search page (to be implemented) will collect information combined in all categories and deliver output in formats ready for analysis programs. Given the volume of information, LAPOGED*b* initiates a new way of organizing and querying data produced by LAPOGE UFSC.

Keywords: Genetic Epidemiology; data organization; genetic markers

LISTA DE FIGURAS

- Figura 1: Escala exemplificando os valores e casas decimais de bytes....21
- Figura 2: Infográfico modificado pelo autor, exemplificando o espaço físico que os dados gerados diariamente ocupariam, caso fossem gravados em CD's.....22
- Figura 3: Estrutura de um Banco de Dados. A estrutura de um BD é dividida em três partes: A primeira (1) consiste de um sistema de gerenciamento de banco de dados; a segunda (2) de um local (servidor), onde são armazenados os dados; a terceira (3) é a interface gráfica para o usuário do banco realizar as consultas e gerar relatórios das pesquisas realizada no banco de dados.....30
- Figura 4: Exemplo hipotético de um Diagrama Entidade Relacional para Banco de Dados. Os retângulos representam entidades, losangos são as relações e os atributos são conectados por linhas contínuas. Neste exemplo, a raiz é a entidade [prontuário] e todas as relações e os atributos são conectados por linhas contínuas. Neste exemplo, a raiz é a entidade [prontuário] e todas as relações e atributos se conectam a ela por meio de outras entidades.....38
- Figura 5: Especificações para o servidor da UFSC para gerenciamento de Banco de Dados.....39
- Figura 6: Pagina inicial do LAPOGE, destacado em vermelho link para acesso ao LAPOGEDb.....42
- Figura 7: Página inicial do LAPOGEDb. Campos de identificação de usuário através de nome e senha e links relacionados ao LAPOGE e estudos de genética epidemiológica.....43
- Figura 8: Aba de dados pessoais do LAPOGEDb. Campos de identificação inicial e seleção de tipos de questionário.....44
- Figura 9: Campos de endereço e dados de contato destacadas como sigilosas do LAPOGEDb.....46
- Figura 10: Campo de informações pessoais do LAPOGEDb.....47

Figura 11: Campo de antecedentes familiares com caixas de seleção do LAPOGEDb.....	48
Figura 12: Campo de hábitos tabagistas na aba de dados epidemiológicos do LAPOGEDb.....	49
Figura 13: Histórico de amamentação do LAPOGEDb.....	50
Figura 14: Histórico diverso de doenças e investigação de exames para doenças infecto-contagiosas do LAPOGEDb.....	51

SUMÁRIO

1 INTRODUÇÃO	19
1.1 EVOLUÇÃO DO ARMAZENAMENTO DE DADOS	19
1.2 <i>BIG DATA</i>	21
1.3 PGH, 1000 GENOME E OUTROS.....	24
1.4 BIOINFORMÁTICA.....	27
1.5 BANCO DE DADOS.....	28
1.6 BANCO DE DADOS BIOLÓGICOS E A IMPORTÂNCIA DA ORGANIZAÇÃO DE DADOS EM LABORATÓRIOS DE GENÉTICA.....	31
1.7 HISTÓRICO DO LAPOGE.....	32
2. JUSTIFICATIVA.....	33
3 OBJETIVOS	35
3.1 OBJETIVO GERAL.....	35
3.1.2 Objetivos Específicos	35
4. MATERIAL E MÉTODOS.....	37
5. RESULTADOS E DISCUSSÃO	41
5.1 LEVANTAMENTO DE DADOS LAPOGE.....	41
5.2 BANCO DE DADOS LAPOGE.....	41
5.2.1 Acesso e definições de usuário.....	42
5.3 PADRONIZAÇÃO DOS QUESTIONÁRIOS INVESTIGATIVOS.....	44
5.3.1 Sessões comuns aos questionários.....	44
5.3.2 Sessões específicas de prontuário.....	51
6. CONSIDERAÇÕES FINAIS	53
Perspectivas futuras:.....	53
REFERÊNCIAS	55

APÊNDICE A – Trabalhos realizados no LAPOGE, organizados por nome do autor	58
APÊNDICE B - Lista de marcadores utilizados nos trabalhos realizados no LAPOGE ao decorrer dos 22 anos de sua existência, organizados por nome do marcador.....	64
APÊNDICE C- Estrutura relacional do LAPOGEDb	72
APÊNDICE D - Campo de dados familiares do LAPOGEDb.....	73
APÊNDICE E - Campo de dados epidemiológicos do LAPOGEDb..	74
APÊNDICE F – Campo de dados biológicos do LAPODEDb	76
APÊNDICE G – Campo de dados médicos do LAPOGEDb	78
APÊNDICE H – Campo de amostra biológica do LAPOGEDb	80
APÊNDICE I – Campo de Artrite do do LAPOGEDb	81
APÊNDICE J – Campo de Câncer de mama do LAPOGEDb	84
APÊNDICE K – Campo de Lupus do LAPOGEDb	85
APÊNDICE L – Campo de psoríase do LAPOGEDb	87

1 INTRODUÇÃO

1.1 EVOLUÇÃO DO ARMAZENAMENTO DE DADOS

O maior passo para o avanço da comunicação humana foi seguramente a revolução da escrita, quando passamos de uma comunicação oral primitiva e interativa direta, para a comunicação através da escrita baseada em textos lineares e uso sistematizado de letras (DIAS, 1999). A escrita trouxe consigo a oportunidade de guardar histórias e informações que normalmente seriam facilmente perdidas e/ou distorcidas, pois seria praticamente impossível garantir que após vários eventos de transmissão a mensagem se manteria fiel a original, uma vez que depende da capacidade humana de memorização, além daquela de emissão e recepção auditiva, estando sujeita ainda a interpretações ou adaptações de tradução diferentes. Essas informações puderam ser passadas adiante não só para as gerações presentes como as futuras, sem grandes perdas ou desvios e em um contexto totalmente diferente do que aquele anteriormente produzido, permitindo a estruturação de ideias muito mais complexas; da mesma forma que anotar números em um papel possibilita a resolução de problemas matemáticos muito maiores.

A introdução da escrita foi um grande avanço de domínio técnico, sendo um sistema eficaz na capacidade de exprimir graficamente a linguagem. Não é para menos que os povos que adquiriram primeiramente esta forma de comunicação textual tiveram o desenvolvimento de suas construções culturais muito acelerados (COSTA, 2002).

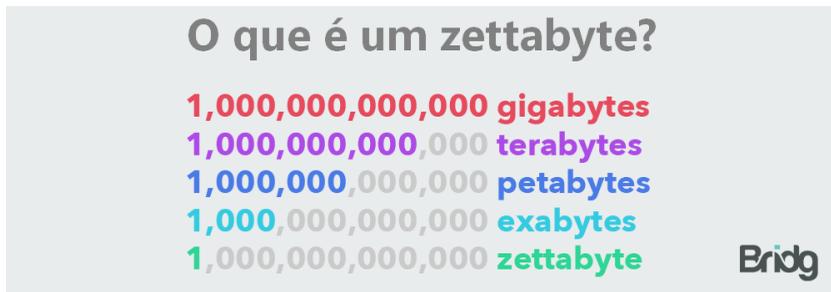
Milhares de anos após o surgimento da escrita, todos os materiais escritos que a humanidade possuía eram ainda produzidos por escribas. Houve uma expressiva mudança de parâmetros no período moderno. A revolução da imprensa começou em 1445, quando Johannes Gutenberg inventou a prensa móvel, uma máquina de impressão tipográfica que poderia fazer em horas, o trabalho de uma vida inteira de um escriba. Com a invenção da prensa móvel até o final do século XV a Europa já havia impresso quase vinte milhões de cópias de livros, e nos séculos subsequentes essas produções passaram para duzentos milhões no século XVI e atingiram um bilhão no XVIII. Essa grande produção e

popularização de livros, fez com que as ideias e novas informações pudessem circular com muito mais facilidade e a ciência pode finalmente discutir com liberdade os seus resultados.

Mas a prensa foi só o começo da nossa grande produção de dados: uma nova grande evolução ocorreu com a invenção de Alan Turing, por volta de 1936. Este matemático inglês foi considerado o pai da computação e tornou-se o precursor dos computadores modernos, ao desenvolver o que foi chamado de a “máquina de Turing”. Ela era capaz de fazer cálculos matemáticos e desempenhou um grande papel no desenvolvimento da ciência da computação e na formalização do conceito de algoritmo. Utilizava uma lógica de código binário, através de “sim ou não”, “Zero e um” ou um bit, e com base nessas simples instruções seguia os algoritmos e executava programas (TURING, 1938). A lógica utilizada na máquina de Turing é essencialmente o cerne da computação. Esta ideia incrivelmente simples que foi encontrada na invenção de Turing é tão poderosa que basicamente todos os computadores e programas modernos a utilizam.

Quando os primeiros programas estavam sendo escritos, o conjunto de oito bits utilizado para codificar uma letra qualquer, foi nomeado de “byte”. Os primeiros computadores inventados trabalhavam com um volume de oito mil bites ou um kilobyte, máquinas como o RAMAC 340 ofereciam até milhões de bytes ou megabytes de armazenamento. De lá pra cá, nosso armazenamento aumentou de capacidade, ficou menor fisicamente e mais barato. Por volta de 2003 passou-se a produzir mais dados digitais armazenados em HD’s, DVD’s e memórias do que em meios analógicos como filmes, fotos, livros e jornais, em pouco tempo a escala inicial de kilobyte dos primeiros computadores passou para os mega, giga, tera e petabytes. Tudo o que a humanidade havia produzido (filmes, livros, series, músicas, fotos jornais entre outros) até 1986 ocupava dois mil e seiscentos petabytes ou 2,6 exabytes, já em 2007 esses valores atingiram duzentos e noventa e cinco exabytes e em 2011 inaugurou-se então uma nova casa decimal, a humanidade já havia acumulado mais de mil e seiscentos exabytes ou 1,6 zettabytes (HILBERT; LÓPEZ, 2011), como exemplificado na Figura 1.

Figura 1: Escala exemplificando os valores e casas decimais de bytes.



Fonte: <http://bridg.com/12-very-shareable-stats-on-big-data/>

Um único canal de vídeos (*Youtube*) apenas no seu primeiro ano de vida possuía um exabyte de dados, o que a primeira vista pode parecer incrível, porém não se compara com a quantidade de informações que a natureza guarda. Por exemplo, o observatório astronômico ASKAP (*Australian Square Kilometre Array Pathfinder*) produz atualmente 7,5 terabytes por segundo de dados amostrais na forma de imagem, no entanto há previsões de que essa taxa aumente em cem vezes até 2025, chegando a gerar sozinho em torno de vinte e cinco zettabytes por ano; com o sequenciamento de genomas cada vez mais barato, calcula-se que em 2025 seja atingido o objetivo de dois bilhões de pessoas sequenciadas (STEPHENS et al., 2015).

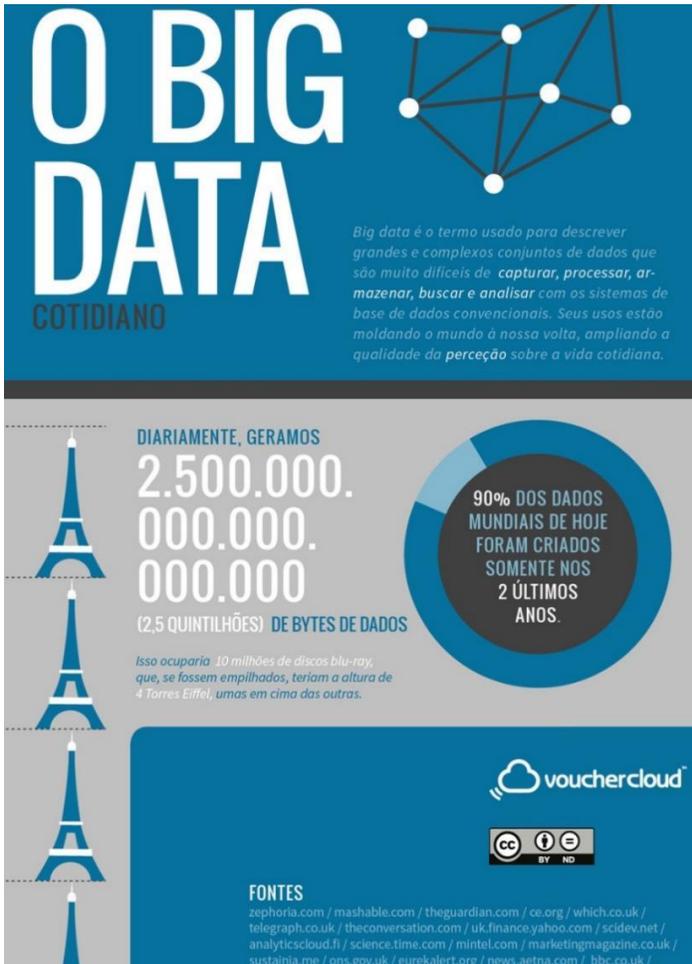
1.2 BIG DATA

Se antes o mais provável era que a maior parte da informação se perdesse, hoje se espera que tudo seja preservado. Isso nos coloca em uma era onde navegamos em um dilúvio de dados, conhecido usualmente sob o termo *big data*. Em tecnologia da informação, *big data* é o termo que descreve o imenso volume de informações e dados acumulados e armazenados.

Diariamente são gerados 2,5 quintilhões de bytes de dados; se essa quantidade de dados fosse gravada em CD's, isso ocuparia cerca de dez milhões de discos Blu-ray, que se empilhados, teriam a altura aproximada de quatro Torres Eiffel, uma em cima das outras). Surpreendentemente 90% dos dados do mundo todo foram gerados nos últimos dois anos (Figura

2). Esse crescimento exponencial está correlacionado principalmente com a adesão das grandes empresas à Internet e criação de redes sociais e dispositivos móveis (HILBERT; LÓPEZ, 2011).

Figura 2: Infográfico modificado pelo autor, exemplificando o espaço físico que os dados gerados diariamente ocupariam, caso fossem gravados em CD's.



Fonte: <http://www.qualimedia.com.br/blog/saiba-como-o-big-data-pode-facilitar-sua-vida/>

O impulso dado pela tecnologia, principalmente pelo incremento do uso dos dispositivos móveis e a profusão de informações na Internet trouxe um forte aumento no volume de dados. Originados pelas mais diversas fontes tais dados ocasionam uma sobrecarga de informações disponíveis para a sociedade, sendo importante registrar que apenas 1% destes dados é efetivamente analisado (JOSÉ; RIBEIRO, 2014).

Os grandes avanços tecnológicos e a disponibilidade cada vez maior de novas ferramentas, torna a geração de dados na ciência cada vez mais simples e acessível aos pesquisadores, o reflexo disso é a quantidade extraordinária de dados que foram produzidos e armazenados ao longo das últimas décadas em nível global.

A aquisição de dados para a genômica envolve formatos muito heterogêneos e é amplamente distribuída, o que torna a taxa de crescimento estonteante, com a quantidade total de sequências genéticas produzidas sendo dobrada a cada sete meses. Os sequenciamentos brutos usados na maioria dos estudos publicados são arquivados na *Sequence Read Archive* (SRA) mantido pelos *United States National Institutes of Health National Center for Biotechnology Information* (NIH/NCBI) ou uma das suas contrapartes internacionais. A SRA atualmente contém mais de 3,6 petabases de dados de sequências brutas, o que inclui cerca de 32.000 genomas microbianos, 5.000 genomas de plantas e animais, e 250.000 sequências de genoma humano individuais já finalizadas ou em andamento. No entanto, os 3,6 petabases* representam uma pequena fração do total produzido; a maior parte dele ainda não está nestes arquivos. Com base nas especificações do fabricante dos instrumentos, pode-se estimar que a capacidade de sequenciamento anual excede os 35 petabases por ano e considerando que o ritmo de crescimento mantenha-se sendo duplicado a cada sete meses, devemos então nos aproximar a um zettabase de sequências por ano até 2025 (STEPHENS et al., 2015).

Essa quantidade exorbitante de dados é maior que a organização e controle destes: a reunião de dados é tão rápida que excede a capacidade de analisar, validar e guardar essas informações e a tendência é que continue aumentando cada vez mais (HAAG, 2011). Isso significa que há ainda muito potencial a se extrair, há ainda muito a se descobrir. Uma

* O valor de conversão aproximado para usar na interpretação destes números é de 4bases = 1byte. (Stephens, 2015)

análise adequada desses grandes conjuntos de dados poderia permitir o encontro de correlações até então não descobertas. Levando em consideração a quantidade impressionante de informações agregadas, apenas uma pequena parcela foi analisada a fim de extrair todas as informações que aquele dado em específico poderia nos fornecer.

A quantidade inverossímil de informações traz grandes limitações na organização sistematizada dos dados, pois sem a utilização de ferramentas de gerenciamento e sem uma boa organização dos mesmos, torna-se praticamente impossível visualizar respostas ou perguntas científicas que poderiam saltar aos olhos comparando com situações onde existe uma boa ferramenta de busca e organização.

1.3 PGH, 1000 GENOME E OUTROS

O projeto genoma humano (PGH) pode ser considerado como uma das maiores façanhas científica da história da humanidade. Foi um projeto de esforço global para sequenciar e mapear todos os genes e regiões intergênicas de humanos. Teve seu início oficial em 1989, com a participação de aproximadamente cinco mil cientistas distribuídos em centenas de laboratórios pelo mundo e tinha como previsão estimada quinze anos até a sua finalização, a partir de um consórcio público liderado pelo pesquisador James Watson, posteriormente por Michael Gottesman e finalmente por Francis Collins (MOREIRA, 2015).

A ideia do consórcio foi fragmentar o DNA alvo, sequenciar os nucleotídeos e ao final, publicar seus resultados parciais em um site de acesso público. Disponibilizar os dados significaria impedir intencionalmente o registro de patente e a retenção sigilosa da informação produzida durante as etapas intermediárias da pesquisa.

Todavia, devido a diversos motivos (tais como divisão de recursos, intenção de patente e disputa pela coordenação do projeto geral), em maio de 1998, o cientista Craig Venter saiu do consórcio público e entrou na corrida pelo sequenciamento do genoma humano, criando uma empresa privada que prometia finalizar os sequenciamentos muito antes do consórcio público. Em decorrência dessa competição, o consórcio público (formado pelo Instituto Nacional de Saúde e o Departamento de Energia – ambos norte-americanos) se viu obrigado a apressar suas pesquisas com o

receio constante de que, caso o setor privado terminasse antes, os genes humanos seriam patenteados (CORRÊA, 2002).

A conclusão do Projeto Genoma Humano acabou por ser anunciada em 26 de junho de 2000, na Casa Branca pelos dois líderes das pesquisas. Sendo atribuídos então a ambos, consórcio público e empresa privada (Celera Genomics), os méritos da pesquisa. Cada um publicou seus resultados em revistas de grande prestígio no meio científico: o consórcio público publicou na revista científica *Nature* e a Celera, laboratório de Craig Venter, publicou seus resultados na revista científica *Science* 24 horas depois. Contrariando as expectativas iniciais, o projeto chegou ao resultado de aproximadamente três bilhões de pares de bases distribuídos em vinte e quatro cromossomos, sendo que apenas 3% do genoma humano foi caracterizado como passível de transcrição (MOREIRA, 2015).

Em sequência, surge em 2008 o 1000 Genome, outro projeto de consórcio público e colaboração internacional. Consistiu em sequenciar o genoma inteiro de mais de mil indivíduos, provenientes de populações diferentes em todo o mundo, com o objetivo central de descrever a maior parte das variantes genéticas que ocorrem com uma frequência superior a 1% na população (VIA et al., 2010). Além destes objetivos científicos principais de criar tanto um catálogo detalhado da variação genética humana e métodos exaustivos para descobrir com precisão e caracterizar a variação usando novas tecnologias de sequenciamento, os membros do centro de coordenação de dados do projeto desenvolveram e implantaram diversas ferramentas que permitem que o acesso aos dados seja público e generalizado (CLARKE et al., 2012). Ao final do projeto, em 2015, foram reconstruídos os genomas de 2.504 indivíduos, provenientes de 26 populações; caracterizando no total 88 milhões de variantes genéticas (84,7 milhões de polimorfismos de nucleotídeo único (SNPs), 3,6 milhões de inserções/deleções (InDels), e 60.000 variantes estruturais)(AUTON et al., 2015).

No Brasil, o primeiro projeto genoma foi iniciado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) em 1998. O projeto proposto na época, foi o sequenciamento da bactéria fitopatogênica *Xylella fastidiosa* (SIMPSON E REINACH et al. 2000). A rede paulista para sequenciamento de *X. fastidiosa* foi batizada de ONSA (sigla em inglês para: Organização para Sequenciamento e Análise de Nucleotídeos),

além de fazer uma analogia clara ao felino brasileiro. A Coordenação do projeto ficou sob os cuidados do britânico Andrew Simpson do Instituto Ludwig de Pesquisa sobre o Câncer e contou com a participação de trinta unidades de pesquisa no estado de São Paulo. A iniciativa da FAPESP foi um sucesso e o genoma de *X. fastidiosa* foi publicado dois anos depois, no ano 2000, sendo a capa da revista *Nature* como o primeiro genoma de um fitopatógeno a ser completamente sequenciado no mundo (MOREIRA, 2015).

O Projeto Genoma *Xylella* contribuiu de maneira significativa para o desenvolvimento da Bioinformática no Brasil. O custo deste projeto foi de aproximadamente 13 milhões de dólares, investidos não apenas em equipamentos de última geração para sequenciamento e análise funcional de genes mas também no suporte e capacitação/treinamento em todos os níveis acadêmicos (graduação, mestrado, doutorado e pós-doutorado). Tais investimentos em infraestrutura e capacitação de pesquisadores foram essenciais para o surgimento, em seguida, de três outros importantes projetos: SUCEST (*Sugar Cane EST*), *Xanthomonas campestris* (segundo fitopatógeno) e do Câncer Humano (HCGP) (MOREIRA, 2015).

Os métodos automatizados de sequenciamentos do genoma e suas aplicações desenvolvidos durante o PGH e os outros projetos genoma subsequentes, tiveram como principal consequência o acelerado desenvolvimento da bioinformática, que suporta de forma essencial as pesquisas atuais realizadas na área da genômica (CORRÊA, 2002). Porém, o grande volume de dados biológicos gerados com a finalização da primeira etapa do PGH evidenciou o que até então era esperado e, de certa forma, temido pelos pesquisadores envolvidos no projeto: o PGH e todas suas consequências estariam fadados a falhar caso a tecnologia computacional (principalmente no campo da sistematização de anotações biológicas e armazenamento de informações) não evoluísse de forma concomitante e progressiva com o projeto (ROBBINS, 1994). Além disso, é importante ressaltar que, por mais avançadas que sejam as técnicas de sequenciamento, elas apenas permitem determinar a ordem dos nucleotídeos no genoma do organismo em questão; o sucesso dos projetos genoma depende não apenas da evolução das tecnologias informacionais, como também da capacitação dos pesquisadores e suas habilidades em interpretar esses dados de sequenciamento e variações genéticas (JORDAN et al., 2010; MOREIRA, 2015).

1.4 BIOINFORMÁTICA

As origens da bioinformática, tanto o termo quanto a disciplina, são difíceis de datar (ATTWOOD et al., 2011). Entretanto, tornou-se rapidamente um campo emergente da pesquisa, que utiliza ferramentas e técnicas computacionais para o armazenamento, análise e apresentação de dados biológicos e moleculares, o que é especialmente importante para análise dos dados sobre pesquisa do genoma (ALVES, 2013). Pode ser definida ainda como a convergência entre as ciências biológicas e ciências da computação, vínculo que possibilita a construção e utilização de ferramentas de software para gerar, armazenar, acessar, processar e analisar dados e informações (STEIN, 2003).

Durante a década de 1960, grupos de pesquisadores das mais diversas áreas das ciências biológicas viram-se diante da necessidade de dedicar esforços em aprender aspectos computacionais para aplicar em seus objetos de estudo. Posteriormente, o início do projeto genoma humano na década de 1990 proporcionou o grande impulso para o desenvolvimento da bioinformática e o seu estabelecimento como disciplina. Além disso, o desenvolvimento e o uso cada vez mais difundido da internet na década de 1990 tornaram possíveis o acesso instantâneo, o intercâmbio e a divulgação de dados biológicos (HIGGINS E TAYLOR, 2001).

Com o crescimento no volume de dados e a evolução das ferramentas e técnicas de bioinformática, tornaram-se cada vez mais frequentes as pesquisas *in silico*, que se referem às experimentações de cunho biológico realizadas inteiramente com ferramentas de bioinformática. Esse tipo de pesquisa tem ocorrido recorrentemente e também tem mostrado cada vez mais sua importância.

É possível perceber que após o advento do projeto genoma humano e outros projetos de sequenciamento de genoma, criou-se um gigantesco e crescente volume de sequências cujas relações evolutivas e funcionais precisam ser elucidadas (VERLI, 2014). Diariamente são geradas grandes quantidades de dados através de pesquisas com abordagens experimentais, e tem se tornado cada vez mais difícil gerenciar e interpretar toda essa abundância de informações. Essas pesquisas não só utilizam dados que de outra forma poderiam se tornar obsoletos como podem perceber padrões ou novas informações cruzando dados de estudos diferentes que, quando

olhados individualmente ou com uma visão restrita, não são capazes de demonstrar todo seu potencial.

Apesar da postulação e da exigência de validação experimental para estudos computacionais, não é raro que um dado computacional apresente maior precisão que um dado obtido na bancada. Os modelos computacionais podem ser ordinariamente chamados de teóricos quando comparados aos modelos ditos experimentais, sendo que na verdade são apenas experimentos computacionais que muitas vezes não possuem uma contraparte em experimentos de bancada e que frequentemente podem complementar ou até mesmo se mostrarem cruciais para a finalização de um estudo iniciado em modelo experimental. O modelo em si e os adjetivos que o acompanham não carregam consigo qualificações quanto a confiabilidade de seus resultados gerados (VERLI, 2014).

1.5 BANCO DE DADOS

Os bancos de dados e sistemas de banco de dados já se tornaram um componente essencial na vida da sociedade moderna, fazemos uso desta ferramenta até mesmo sem perceber em diversas atividades do nosso cotidiano. Como por exemplo, ao fazer compras *online* ou consultar algum livro em uma biblioteca virtual; ao utilizar um estabelecimento bancário para depositar ou retirar fundos e até mesmo a compra de produtos em supermercados atualiza automaticamente o banco de dados que mantém o controle de estoque dos itens. Todos esses são exemplos de aplicações tradicionais de banco de dados, nos quais a maior parte da informação armazenada e acessada é textual ou numérica. Nos últimos anos os avanços na tecnologia fizeram com que fosse possível gerar novas aplicações para o sistema de banco de dados, sendo possível a utilização de armazenamento de imagens, clipes de áudio e *streaming* de vídeos digitalmente. Este último formato de arquivo se tornou muito importante para os sistemas de informações geográficas (GIS – *Geographic Information Systems*) que podem armazenar e analisar mapas, dados sobre o clima e imagens de satélite (ELMASRI E NAVATHE, 2010).

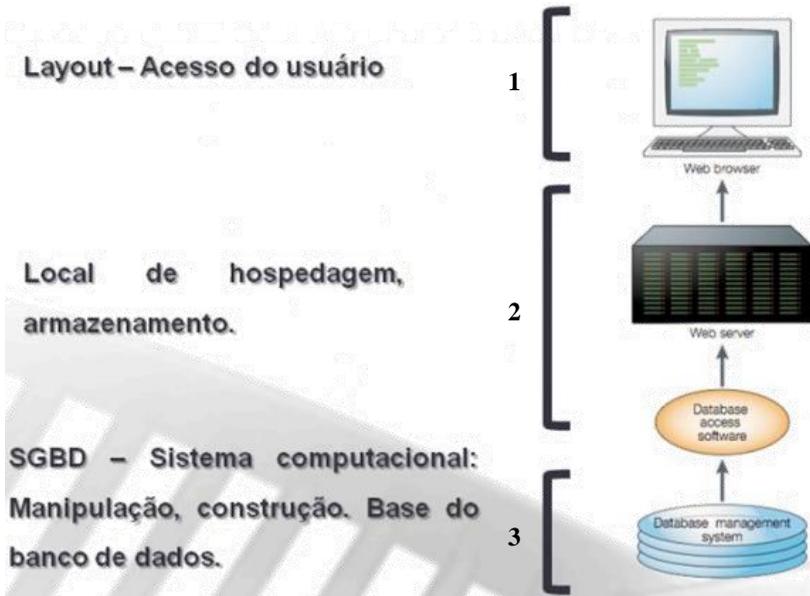
Um Banco de Dados (BD) é um arquivo informatizado utilizado para armazenar e organizar os dados de tal forma que tais informações podem ser recuperadas facilmente através de uma variedade de critérios de pesquisa. O objetivo principal do desenvolvimento de um BD é organizar dados em um conjunto de registros estruturados para permitir a fácil

recuperação de informações. Cada registro, também chamado de *entry*, deve conter um número de campos que possuem os itens de dados reais, por exemplo, campos para nomes, números de telefone, endereços, datas, entre outros. Para recuperar um determinado registro do BD, o usuário pode especificar uma determinada informação, chamada de *value*, para ser encontrado em um campo particular e o computador então recupera todo o registro de dados relacionados à esta informação. Este processo é chamado de *making a query* (do inglês, fazendo uma consulta) (WILTON E COLBY, 2005).

De forma geral, laboratórios de pesquisa de pequeno porte utilizam arquivos comuns ou planilhas (por exemplo, Excel®) como BD. Este tipo de abordagem é útil para armazenar pequenas quantidades de informações e exige pouco poder de trabalho para um computador, entretanto o número de informações possíveis de serem armazenadas é finito. À medida que o tamanho do BD aumenta ou se os tipos de dados se tornam mais complexos, o estilo de BD através de arquivos comuns ou planilhas pode tornar a recuperação da informação uma tarefa árdua, frequentemente causando falhas no sistema devido à falta de memória ou processamento para executar a tarefa (DATE, 2003).

Para facilitar o acesso e recuperação de dados, diversos e sofisticados programas de computador têm sido desenvolvidos para organizar, pesquisar, processar e acessar dados. Eles são chamados de Sistemas de Gerenciamento de Banco de Dados (SGBD). Esses sistemas contêm não somente os registros de dados, mas também as instruções operacionais para ajudar a identificar as conexões entre os registros, favorecendo os processos de definição, construção, manipulação e compartilhamento de dados entre vários usuários e aplicações. O objetivo da criação de uma estrutura de dados por um SGBD é facilitar a execução de pesquisas e combinar diferentes registros para formar relatórios de pesquisa final. Um dos modelos de SGBD mais utilizados é o modelo de BDR - Banco de Dados Relacionais (ELMASRI E NAVATHE, 2010). A figura 3 demonstra de forma simplificada a arquitetura de um BD padrão com o uso de um SGBD (STEIN, 2003).

Figura 3: Estrutura de um Banco de Dados. A estrutura de um BD é dividida em três partes: A primeira (1) consiste de um sistema de gerenciamento de banco de dados; a segunda (2) de um local (servidor), onde são armazenados os dados; a terceira (3) é a interface gráfica para o usuário do banco realizar as consultas e gerar relatórios das pesquisas realizadas no banco de dados.



Fonte: STEIN, 2003.

Ao invés de usar uma única tabela como em uma planilha, BDRs usam um conjunto de tabelas para organizar os dados. Cada tabela, também chamada de “relação”, é composta de linhas e colunas. As colunas representam os campos individuais. As linhas representam os valores nos campos de registros. As colunas de uma tabela são indexadas de acordo com uma característica comum chamado “atributo”, para que eles possam ter uma referência cruzada em outras tabelas. Para executar uma consulta em um BDR, o sistema seleciona os itens de dados ligados de diferentes tabelas e combina as informações em um único relatório. Portanto, informações específicas podem ser encontradas mais rapidamente a partir de um BDR, do que a partir de um BD de arquivo simples ou de uma única tabela (STEIN, 2013).

Os BDR podem ser criados usando uma linguagem de programação chamada *Structured Query Language* (SQL). A criação deste tipo de bases de dados é trabalhosa e depende de um grande planejamento durante a fase de construção e modelagem do banco. Após a criação da base de dados uma nova categoria pode ser facilmente adicionada sem exigir que todas as tabelas existentes tenham que ser também modificadas. Fatores tais como código aberto, alta velocidade de processamento e disponibilidade para vários sistemas operacionais, devem ser levados em consideração para escolher o SGBD a ser utilizado na construção do BD. O SGBD MySQL utiliza a linguagem SQL e dada sua rapidez, capacidade de manusear grandes quantidades de dados e por atender todas as especificações citadas acima, o MySQL pode ser considerado como um dos BDR de maior sucesso no mundo atual (TAHAGHOGHI E WILLIAMS, 2007).

1.6 BANCO DE DADOS BIOLÓGICOS E A IMPORTÂNCIA DA ORGANIZAÇÃO DE DADOS EM LABORATÓRIOS DE GENÉTICA

Bancos de Dados Biológicos (BDB) são considerados essenciais para a experimentação e análise biológica. Como tal, eles representam uma importante base nas descobertas científicas (ELTABAKH, 2006). O armazenamento adequado e organizado de dados biológicos é um dos eixos principais da bioinformática, sobretudo no que diz respeito aos BDBs. Eles oferecem aos cientistas a oportunidade de acessar uma grande variedade de dados biologicamente relevantes, incluindo as sequências genômicas de uma gama cada vez mais ampla de organismos (BAXEVANIS, 2011).

O crescente avanço tecnológico e a quantidade de sequenciadores de nova geração têm causado um crescimento no número de genomas sequenciados numa proporção inimaginável. Até dezembro de 2015, foram registrados 36.052 projetos de sequenciamento finalizados (7.747 concluídos e 28.305 classificados como permanentes), 33.571 projetos em andamento e 1.856 projetos de sequenciamento de regiões de interesse do genoma (<http://genomesonline.org>). O número total de projetos, 71.479, evidencia que estamos vivendo uma revolução genômica (MOREIRA, 2015). A enorme quantidade de dados genômicos, provenientes dos diversos projetos de sequenciamento, tornaram os BDBs componentes essenciais no cotidiano da vida científica atual.

Grandes BDB com informações gerais de diversos organismos de acesso fácil e público, como o EMBL (<http://www.ebi.ac.uk/embl/>), DDBJ (<http://www.ddbj.nig.ac.jp/>), NCBI (<http://www.ncbi.nlm.nih.gov/>) e o EBI (<http://www.ebi.ac.uk/>) são considerados referências para modelos de BDB. A integração de dados é uma das funcionalidades mais marcante destes, garantindo que não haja, por exemplo, a sobreposição de informações disponíveis. Isto permite a criação e implementação de bancos menores e especializados, que possam atender as necessidades de cada grupo de pesquisa (STEIN, 2003; MOREIRA, 2015).

1.7 HISTÓRICO DO LAPOGE

O Laboratório de Polimorfismos Genéticos (LAPOGE) foi fundado em 23 de junho de 1994 nas antigas dependências do departamento de Biologia Celular, Embriologia e Genética (BEG), do Centro de Ciências Biológicas da UFSC, tendo como supervisora a Prof. Dra. Ilíada Rainha de Souza.

As atividades do laboratório iniciaram-se através de estudos sobre a variabilidade genética de populações isoladas do estado do Paraná e de Santa Catarina. Em 1994, identificou-se a variabilidade genética de sistemas sanguíneos, proteicos e enzimáticos na comunidade de Witmarsum (PR). Em 1995, a pesquisadora e sua equipe retornaram à comunidade isolada de Valongo, para aquisição de novos dados sobre a população e estudaram-se alguns parâmetros demográficos em que detectou-se a variabilidade genética de alguns sistemas sanguíneos, proteicos e enzimáticos, através de técnicas sorológicas em tubos de ensaio e de eletroforese em gel de amido. De 1995 a 2003 foram feitas coleta nas comunidades da Costa da Lagoa e de São João do Rio Vermelho, para a identificação da variabilidade e similaridade genética destas populações, localizadas na Ilha de Santa Catarina, através de marcadores proteicos e enzimáticos. Entre 1996 e 2000 estudaram-se os polimorfismos dos genes HLA, nestas duas comunidades. Em 2001, a equipe do LAPOGE realizou um estudo da variabilidade proteica e enzimática numa amostra de doadores de sangue da população da grande Florianópolis. No ano seguinte foram implantados estudos de microssatélites e em 2003, com projetos aprovados pelas agências de fomentos, iniciaram-se as pesquisas sobre a

identificação de polimorfismos de genes de baixa penetrância em pacientes com câncer de mama e dengue.

Atualmente o LAPOGE desenvolve projetos epidemiológicos e genéticos, com ênfase especial em Doenças Autoimunes, Câncer de Mama e Estratificação Populacional; projetos de pesquisa acadêmica e de extensão à comunidade, além de abrir espaço para estudantes, formando pesquisadores de graduação e pós graduação.

2. JUSTIFICATIVA

O grande volume de dados gerados pela genotipagem de marcadores durante os 22 anos de LAPOGE, em estudos de epidemiologia genética, cria a necessidade da construção e implementação de um BDB informacional e relacional, pois garante a organização adequada dos dados obtidos, bem como otimiza as buscas e é de grande valia no auxílio para análises.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Construir e implementar, através de ferramentas computacionais, um Banco de Dados Biológicos informacional e relacional para armazenamento de informações biológicas, genéticas, familiares e epidemiológicas coletadas no Laboratório de Polimorfismos Genéticos da Universidade Federal de Santa Catarina (LAPOGE - UFSC), visando organizar, padronizar e diminuir as chances de associações espúrias relacionadas a manuseio incorreto das informações, auxiliando nas análises e metanálises que venham a ser feitas posteriormente.

3.1.2 Objetivos Específicos

- ✓ Realizar levantamento dos trabalhos produzidos pelo LAPOGE, contabilizando marcadores genéticos e informações epidemiológicas utilizadas;
- ✓ Construir BDB utilizando linguagens de programação de alto nível para o armazenamento das informações genéticas, clínicas e epidemiológicas;
- ✓ Implementar este database em um servidor de acesso restrito a integrantes cadastrados no LAPOGE - UFSC;
- ✓ Retirar todas as informações dos questionários aplicados no momento da coleta de dados biológicos visando unificar e padronizar os mesmos;
- ✓ Padronizar o sistema de códigos de identificação que garantem o anonimato dos doadores, mantendo todas as suas informações sigilosas.
- ✓ Desenvolver as linhas de código apropriadas para adaptar o acesso ao banco de dados do LAPOGE através de Tablets, diminuindo os riscos de erros no momento da inserção de dados.

4. MATERIAL E MÉTODOS

Os dados de cada amostra já recolhida e utilizada no LAPOGE, são obtidos através de questionários (dados físicos) que estão catalogados por doença relacionada e/ou população específica, incluindo seus respectivos questionários e os Termos de Consentimento Livre e Esclarecido (TCLEs). Dentre estes, se encontram questionários de doenças autoimunes como Psoríase (PSR), Lúpus Eritematoso Sistêmico (LES) e Artrite Reumatoide (ART), além de questionários referentes aos pacientes diagnosticados com Câncer de Mama (CMA) e, por fim, questionários de populações semi-isoladas e híbrida atual (que não possuem caso clínico conhecido) do estado de Santa Catarina. Os dados virtuais estão armazenados em tabelas (planilhas), assim como os questionários físicos; cada planilha representa um BD referente aos pacientes, controles e populações semi-isoladas.

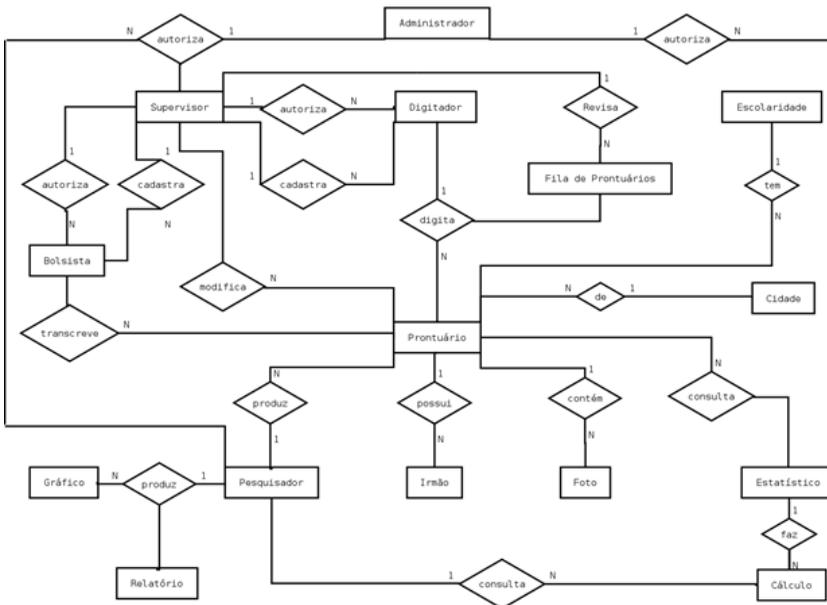
Para a diagramação e programação do BDB, três passos essenciais foram seguidos:

- 1) Levantamento dos trabalhos realizados durante toda a história do laboratório: Tipo de trabalho, ano em que foi desenvolvido e marcador utilizado;
- 2) Identificação das informações (perguntas) comuns em todos os formulários utilizados pelo LAPOGE para a coleta de informações de pacientes, controles e populações em geral;
- 3) Divisão das informações identificadas em categorias: Dados pessoais, dados familiares, dados epidemiológicos, dados biológicos, dados médicos, amostra biológica, marcadores genéticos e informações específicas de CMA, LES, PSR e ART.

Após a identificação e divisão das categorias, empregou-se o Modelo entidade-relacionamento (ER) para a montagem do banco. ER é um modelo conceitual de aplicações de um BD, cuja notação diagramática associada é conhecida como Diagrama ER. O resultado é o registro conciso dos requisitos do usuário, utilizando um diagrama de fluxo de dados que descreve entidades, relacionamentos e atributos, representados por retângulos, losangos e elipses, respectivamente (Elmasri e Navathe, 2008). A figura 04 apresenta o um diagrama ER básico com suas partes essenciais.

Figura 4: Exemplo hipotético de um Diagrama Entidade Relacional para Banco de Dados. Os retângulos representam entidades, losangos são as relações e os atributos são conectados por linhas contínuas. Neste exemplo, a raiz é a entidade [prontuário] e todas as relações e os atributos são conectados por linhas contínuas. Neste exemplo, a raiz é a entidade [prontuário] e todas as relações e atributos se conectam a ela por meio de outras entidades.

Modelagem Conceitual do Banco de Dados (Projeto Febem) Entidade Relacional (ER)



Fonte: (ELMASRI E NAVATHE, 2010).

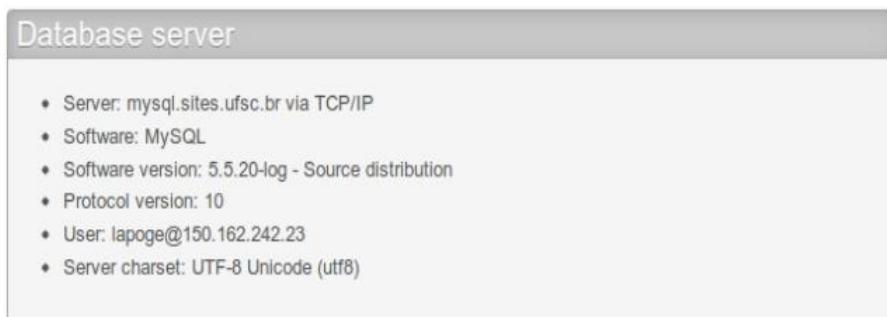
Definido o modelo de ER, o BDB foi programado em linguagem SQL (MySQL), tendo sua interface gráfica desenvolvida em linguagem HTML, com implementações em PHP estabelecendo parâmetros de conexão entre os formulários e MySQL. HTML é a abreviação para *HyperText Markup Language* ou Linguagem de Marcação HiperTexto, uma linguagem simples composta de marcações de formatação e diagramação de hipertexto/hipermídia (informações em texto, imagens, sons e ações ligadas umas às outras) e que se destina a escrever documentos que possam ser lidos por *browsers* (navegadores) (SILVA, 2008). A linguagem de programação PHP (*Hypertext Preprocessor*) é de ampla

utilização, interpretada e de livre acesso, que é especialmente interessante para desenvolvimento de *websites*, podendo ser mesclada dentro do código HTML. Além disso, possui compatibilidade com a linguagem de BD MySQL para criar páginas web dinâmicas que interagem com os visitantes. Com a combinação PHP/ MySQL, pode-se coletar dados de seus usuários e criar conteúdos específicos em tempo real.

A programação foi realizada com o emprego do editor de texto Sublime Text 2, disponível de forma livre (<http://www.sublimetext.com/>). A hospedagem local do BDB foi iniciada em um computador desktop localizado no LAPOGE (Servidor Eukarya - sala 304B). As correções de erros, atualizações e aprimoramentos do BDB foram realizadas com o emprego do phpMyAdmin (uma ferramenta de software livre escrito em PHP, destinado a lidar com a administração do MySQL pela internet ou localmente) (www.phpmyadmin.net).

A hospedagem do BDB foi conectada a um servidor próprio da UFSC (Figura 5), com senha criptografada e intransferível, para a seguridade dos dados, sendo realizadas cópias de segurança três vezes ao dia para garantir a recuperação dos dados armazenados. Eventuais alterações de correções ou *updates* foram primeiramente realizadas localmente no computador desktop do LAPOGE e posteriormente enviadas ao Servidor da UFSC com o auxílio do software livre FileZilla (www.filezilla-project.org).

Figura 5: Especificações para o servidor da UFSC para gerenciamento de Banco de Dados.



Fonte: Autor.

5. RESULTADOS E DISCUSSÃO

5.1 LEVANTAMENTO DE DADOS LAPOGE

Dados provenientes dos 22 anos de existência do LAPOGE foram compilados e indicaram 49 produções sendo 34 trabalhos de conclusão de curso de graduação, 12 dissertações de mestrado e 3 teses de doutorado, que utilizaram um total de 158 marcadores, principalmente em amostras de pacientes diagnosticados com Câncer de Mama e doenças autoimunes como, Psoríase, Lúpus Eritematoso Sistêmico e Artrite Reumatoide, além de estudos em populações semi-isoladas e híbrida atual (que não possuem caso clínico conhecido) do estado de Santa Catarina (Apêndices A e B).

5.2 BANCO DE DADOS DO LAPOGE

O BDB apresentado neste trabalho, foi inicialmente proposto e nomeado de LAPOGEDb por Debortoli (2013). A estrutura relacional criada para conter os dados está diagramada no Apêndice C, o qual mostra a relação de chave primária onde todas as amostras recebem uma identificação única (ID *primary key*), à qual as demais informações são relacionadas. Esta estratégia é usada pois nenhuma informação é isenta de não repetições, isto é, dois indivíduos podem ter nascido no mesmo dia, ter o mesmo endereço, mesmo nome ou ainda o mesmo nome dos pais. A criação de uma chave única primária elimina a possibilidade de confusão nos sistemas de busca (*Query*). O site encontra-se hospedado no domínio [ufsc.br], sendo a solicitação para a utilização do domínio realizada pela Dra. Ilíada Rainha de Souza junto a Superintendência de Governança Eletrônica e Tecnologia da Informação e Comunicação (SETIC) da UFSC (DEBORTOLI, 2013).

Os *scripts* de programação feitos e modificados não foram incluídos nos Anexos ou Apêndices deste trabalho com a intenção de preservar a seguridade do banco e sigilo dos dados. Todavia, podem ser acessados mediante requerimento através do email andrea.marrero@ufsc.br.

5.2.1 Acesso e definições de usuário

O LAPODE *database*, pode ser acessado através de link na página inicial do laboratório (www.lapoge.ufsc.br), como destacado na figura 6, ou diretamente através do endereço eletrônico www.lapoge.sites.ufsc.br.

Figura 6: Pagina inicial do LAPOGE, destacado em vermelho link para acesso ao LAPOGEDb.

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Estadão.ufsc.br Professores.ufsc.br STAF.ufsc.br Comunidade.ufsc.br Extensão.ufsc.br Geral

UFSC - Laboratório de Polimorfismos Genéticos

Laboratório de Polimorfismos Genéticos

Página do LAPOGE

Resultado editais de extensão - 2016

Nosso laboratório foi contemplado, no edital de 2016, em três projetos:

PRODOLSAID:

Vera Costa Netto Muniz	Câncer de mama: Conhecendo, Prevenindo e Evitando Recidivas VII
Andrea Rita Marinho	Curso por EAD: Identificação Molecular Diversidade de peixados comercializados em Santa Catarina v.2
Eliada Ruzita de Souza	Acompanhamento farmacogenético do tratamento de pacientes com Artrose Reumatóide e Síndrome Entenasto (Estudo avaliado no RCV)

Área de saúde:

QR CODE LAPOGE

Links importantes

- BEQ
- CCB
- LAPOGE no Facebook
- LAPOGEDb

Programa de Pós-graduação

Publicações

Contatos

UFSC
Centro de Ciências Biológicas
Campus São Carlos/Paraná/Lima
Sala 7 - Bloco B - 3ª andar - Sala 305
Laboratório de Polimorfismos Genéticos (LAPOGE)
CEP 89040-900
Tafóvia +55 (48) 3721-9804

Fonte: Autor

Na pagina inicial do LAPOGEDb, como pode ser visualizado na figura 7, encontram-se os campos de *login* e senha, além de *links* considerados importantes para a equipe ou pesquisadores relacionados. Os *logins* e senhas foram criados e definidos considerando todos os participantes ativos do laboratório (Professores, pós-doutores, alunos de doutorado, mestrado e graduação), cuja revogação de acesso ocorrerá sempre ao final de um projeto de pesquisa ou caso seja rompido o vínculo dos mesmos com o LAPOGE.

Figura 7: Página inicial do LAPOGEDb. Campos de identificação de usuário através de nome e senha e links relacionados ao LAPOGE e estudos de genética epidemiológica.



Fonte: Autor

O emprego de um sistema de *login* com usuário e senha é essencial para manter a segurança dos dados. Além de certificar que o acesso aos dados seja exclusivo de integrantes ativos do LAPOGE, também se certifica, através de restrições de acesso, quais usuários terão permissão para adicionar ou modificar informações no banco. Ademais, apenas um usuário possui o controle geral do banco, sendo este definido como “Administrador do Banco de Dados” (DBA, do inglês *Database Administrator*). Este usuário é autorizado a realizar alterações na estrutura do banco, podendo: incluir, alterar ou remover categorias dos questionários.

Importante também ressaltar que o controle de acesso ao LAPOGEDb é verificado utilizando um sistema de rastreamento através de data e horário, que registra o último usuário que realizou edições ou atualizações (ELMASRI & NAVATHE, 2010).

5.3 PADRONIZAÇÃO DOS QUESTIONÁRIOS INVESTIGATIVOS

Devido à grande quantidade de questionários distintos utilizados no laboratório ao longo dos anos, sentiu-se a necessidade de padronização para adequamento ao bando de dados. Por isso, foi criado um questionário universal, com perguntas comuns a todos formulários anteriores e adicionadas novas categorias para informações específicas de cada tipo de questionário.

5.3.1 Sessões comuns aos questionários

Ao realizar o *login* no DB, o usuário é direcionado à página de cadastro de formulário. Quando selecionada a opção de cadastrar um novo formulário, ele é então, redirecionado para a primeira aba de preenchimento, denominada Dados Pessoais, onde encontra-se o campo de “Identificação inicial” (figura 8).

Figura 8: Aba de dados pessoais do LAPOGEDb. Campos de identificação inicial e seleção de tipos de questionário.

« voltar Home		Dados Pessoais	Dados Familiais	Dados Genéticos	Informações	Confirmar
IDENTIFICAÇÃO INICIAL						
CÓDIGO	-- seleccione --	→	-- seleccione --			
CODIGO QUESTIONARIO						
Amostra Fantasma?	<input type="radio"/> Sim <input type="radio"/> Nao					
Data desta Entrevista						
Entrevistador	Não Identificado					

Fonte: Autor

Neste momento, o usuário escolhe que tipo de questionário deseja criar, através da opção “Código”. Os códigos são exclusivos e intransferíveis, portanto, são a chave primária de acesso para as informações cadastradas de cada participante. São gerados automaticamente pelo banco utilizando a padronização de três letras e quatro números, estas letras referem-se e são determinadas através do nome de doenças ou populações estudadas no LAPOGE (por exemplo, PSR0001

para identificar a primeira amostra a ser cadastrada em Psoríase). A implementação deste código padronizado e gerado pelo banco, fez-se necessária, uma vez que ao longo destes 22 anos de funcionamento, o laboratório acumulou questionários de uma mesma doença ou população com mais de um tipo de nomenclatura, como por exemplo: CMA, CAMA e CaMa para designar amostras de participantes diagnosticados com câncer de mama. E ao utilizar diferentes nomenclaturas para identificar uma mesma variável, pode-se gerar erros de conflito na busca pela variável em um BD (seja ele construído em tabelas simples ou relacional com o auxílio de um SGBD).

Quando um questionário físico tem seus dados transferidos para o banco, o campo “Código do questionário” é preenchido com o código antigo determinado para o formulário em questão. Este procedimento é importante, pois é necessário que o banco gere um novo código para manter a padronização, porém, também é imprescindível que o código antigo seja mantido e arquivado afim de que, os diversos trabalhos gerados no laboratório (apontados no levantamento inicial) que utilizaram esses códigos anteriores, não perdessem suas identidades amostrais. A “Amostra fantasma” é outro artifício criado para que o código do questionário e o código gerado pelo banco coincidam, visto que existem intervalos numéricos que não correspondem a nenhuma amostra biológica coletada.

Existem nesta sessão ainda, os campos de “Data desta entrevista”, para determinar a idade do participante no momento da entrevista, e identificação do “Entrevistador”.

Subsequente, encontra-se o campo de “Endereço” e “Dados de contato” (Figura 9). Estas informações estão destacadas como “Manter em sigilo”, pois após preenchimento, serão acessíveis apenas para os administradores do banco, esta determinação é indispensável para manter a identidade do participante protegida.

Figura 9: Campos de endereço e dados de contato destacadas como sigilosas do LAPOGEDb.

ENDEREÇO (MANTER EM SIGILO)	
Nome do Entrevistado	<input type="text"/>
CEP	<input type="text"/> Não sabe o CEP? PESQUISE AQUI
Logradouro	<input type="text"/>
Numero	<input type="text"/>
Complemento	<input type="text"/>
Bairro	<input type="text"/>
Cidade/Estado	-- Primeiro seleccione o estado -- <input type="text"/> -- seleccione -- <input type="text"/>

DADOS DE CONTATO (MANTER EM SIGILO)	
Telefone Residencial	<input type="text"/>
Telefone Comercial	<input type="text"/>
Telefone Celular	<input type="text"/>
E-mail	<input type="text"/>
Horario de contato	de <input type="text"/> até <input type="text"/> (usar padrão de 24 horas)
Preferir contato por	<input type="checkbox"/> telefone comercial <input type="checkbox"/> telefone residencial <input type="checkbox"/> telefone celular <input type="checkbox"/> e-mail
Observação	<input type="text"/>

Fonte: Autor

Finalizando a aba de dados pessoais, temos os campos de “Informações pessoais” (figura 10), onde, dentre outras informações, o participante indicará sua profissão. Este dado é relevante em estudos epidemiológicos, pois, a profissão exercida pode estar relacionada com um fator desencadeante de alguma alteração genética ou corporal, sendo assim uma variável a ser considerada durante as pesquisas. Apresenta-se também, o campo de “Etnicidade”, sendo determinada tanto por como o participante se classifica, quanto pelo entrevistador. Esta informação é interessante não apenas para precisar com maior confiança a cor de pele, como para comparar a diferença das respostas em si, podendo indicar discrepância entre como o sujeito se percebe contra como é percebido pela sociedade. Ao final, existe um campo de “Observações”, para inclusão de qualquer informação adicional e os botões de salvar ou limpar as informações preenchidas.

Figura 10: Campo de informações pessoais do LAPOGEDb.

INFORMAÇÕES PESSOAIS	
Genero	<input type="radio"/> Feminino <input checked="" type="radio"/> Masculino
Data de Nascimento*	<input type="text"/>
Local de Nascimento	-- Primeiro seleçõe o estado -- -- seleçõe --
Peso	<input type="text"/> (em quilos, com uma casa decimal)
Altura	<input type="text"/> (com duas casas decimais)
Tipo Sanguineo	<input type="radio"/> A <input type="radio"/> B <input type="radio"/> AB <input type="radio"/> O
Fator Rh	<input type="radio"/> Positivo <input checked="" type="radio"/> Negativo
Profissõe	-- seleçõe -- <input type="checkbox"/> Aposentado(a)
ETNICIDADE (pele participante)	<input type="radio"/> Branco <input type="radio"/> Indio <input type="radio"/> Preto <input type="radio"/> Pardo <input type="radio"/> Amarelo
ETNICIDADE (pele pesquisador)	<input type="radio"/> Branco <input type="radio"/> Mulato <input type="radio"/> Amarelo <input type="radio"/> Negro
ESTADO CIVIL	<input type="radio"/> Solteiro <input type="radio"/> Casado/ União Estável <input type="radio"/> Divorciado <input type="radio"/> Viúvo
ESCOLARIDADE MAXIMA	<input type="radio"/> Analfabeto <input type="radio"/> Ensino Fundamental <input type="radio"/> Ensino Médio <input type="radio"/> Graduação <input type="radio"/> Especialização <input type="radio"/> Pós-Graduação <input type="radio"/> Completo <input type="radio"/> Incompleto
Observaçõe	<input type="text"/>

Fonte: Autor

Ao salvar, o usuário é direcionado para aba de “Dados familiares” (Apêndice D). Esta sessão e todas as subseqüentes apresentarão na parte superior a identificação com o código do formulário gerado automaticamente pelo banco. Nos dados familiares, são adicionadas todas as informações de origem familiar (fraterna, paterna e materna) e históricos de doenças (autoimunes, câncer e outras que forem consideradas relevantes), cujas identidades também são mantidas sob sigilo. O campo de antecedentes familiares possui caixas de seleções nominadas “+ um parente” (figura 11), esta função permite que inúmeros casos de doenças familiares sejam cadastrados.

Figura 11: Campo de antecedentes familiares com caixas de seleção do LAPOGEDb.

Descrição familiar	-- seleccione o parente --	-- seleccione o tipo --	+ um parente >>
Observação			
Histórico doença autoimune?	<input type="radio"/> Não <input type="radio"/> Sim		
Descrição familiar	-- seleccione o parente --	-- seleccione o tipo --	+ um parente >>
Observação			
Histórico de outras doenças?	<input type="radio"/> Não <input type="radio"/> Sim		
Descrição familiar	-- seleccione o parente --	-- seleccione o tipo --	+ um parente >>
Observação			

Fonte: Autor

O cadastro dos antecedentes familiares das amostras é extremamente relevante para a determinação do componente genético das variáveis epidemiológicas.

Após os dados familiares salvos, existe a aba de “Dados epidemiológicos”, nesta é realizado um levantamento dos hábitos relacionados a uso de drogas lícitas e ilícitas e a prática de exercícios. (Apêndice E). Há uma categoria, especialmente para o levantamento de hábitos tabagistas (figura 12), que foi adicionada visto que diversos estudos apontam o tabaco como um fator determinante no desenvolvimento de algumas doenças, por tanto é sempre pertinente saber se o participante fuma ou não, se já fumou e parou, por quanto tempo, que tipo, com que frequência e se convive com alguém que fuma.

Figura 12: Campo de hábitos tabagistas na aba de dados epidemiológicos do LAPOGEDb.

HÁBITOS TABAGISTAS	
Você fuma?	<input type="radio"/> Não, nunca fumei <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, fumo atualmente
Que tipo de cigarro?	<input type="checkbox"/> Cigarro comercial <input type="checkbox"/> Cigarro caseiro <input type="checkbox"/> Charuto <input type="checkbox"/> Cachimbo <input type="checkbox"/> Outros
Qual a quantidade?	<input type="checkbox"/> Um cigarro <input type="checkbox"/> 2 a 3 cigarros <input type="checkbox"/> 3 a 5 cigarros <input type="checkbox"/> mais que 5 cigarros
Com que frequência?	<input type="checkbox"/> Esporadicamente <input type="checkbox"/> Uma vez por dia <input type="checkbox"/> 2 a 5 vezes por dia <input type="checkbox"/> 5 a 10 vezes por dia <input type="checkbox"/> Mais que 10 vezes por dia
Convive com alguém que fuma?	<input type="radio"/> Não <input type="radio"/> Sim
Como você classifica o hábito tabagista desta pessoa?	<input type="radio"/> leve <input type="radio"/> moderado <input type="radio"/> acentuado <input type="radio"/> pesado
Observação	

Fonte: Autor

A próxima aba é a de “Dados biológicos” (Apêndice F) e contempla as informações de alergias, histórico hormonal, uso de contraceptivos, paridade, histórico de abortos e histórico de amamentação. Durante esta etapa é possível perceber que há duas opções para as respostas negativas “não” e “não se aplica por ser homem”, esse artifício foi escolhido para que o banco pudesse reconhecer se um determinado participante não havia amamentado de fato ou se esse critério não poderia ser aplicado a ele, por ser um indivíduo do sexo masculino. Há outras opções ainda, como pode ser visto na figura 13, para investigar os outros porquês das respostas negativas. Pois por exemplo, uma mulher pode ter escolhido não amamentar seus filhos ou ter sido impedida disso por causa do trabalho, mas também pode ter sido incapaz de amamentá-los por limitações biológicas. Saber o motivo pode fazer toda a diferença nos estudos epidemiológicos realizados no laboratório.

Figura 13: Histórico de amamentação do LAPOGEDb.

HISTÓRICO DE AMAMENTAÇÃO	
Você amamentou?	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não amamentei pois não tive filhos <input type="radio"/> Não amamentei meus filhos por decisão própria <input type="radio"/> Não amamentei meus filhos por causa do trabalho <input type="radio"/> Não amamentei meus filhos por limitações biológicas
Quanto tempo?	<input type="text"/>
Observações	<input type="text"/>

Fonte: Autor

Na sessão de “Dados médicos” (Apêndice G), ficam registrados todos os históricos médicos possíveis de conseguir do entrevistado. Esta categoria, além de dar pistas para entendimento dos resultados gerados e tornar possível as conexões de informações de histórico médico com os dados genéticos encontrados nas pesquisas, também é extremamente importante para a segurança do pesquisador que manuseará as amostras. Como mostrado na figura 14, existe um campo nesta aba que abarca doenças diversas, inclusive doenças que apresentam um alto risco biológico durante a manipulação da amostra. Onde mais especificamente também é indagado ao indivíduo se não possui doença infecto-contagiosa, e se declara que não possui porque já realizou exames médicos para determinar essa informação ou se pressupõe que não, porém nunca realizou exames investigativos.

Figura 14: Histórico diverso de doenças e investigação de exames para doenças infecto-contagiosas do LAPOGEDb.

HISTÓRICO DIVERSO	
Algum relato neurológico?	<input type="checkbox"/> Asma <input type="checkbox"/> Dengue <input type="checkbox"/> Hipercolesterolemia <input type="checkbox"/> Hipertrigliceridemia <input type="checkbox"/> Hepatite <input type="checkbox"/> Aids <input type="checkbox"/> Osteoporose <input type="checkbox"/> Tuberculose <input type="checkbox"/> Outro (descrever nas observações)
Informar todos os medicamentos que utiliza atualmente	<input type="text"/>
Observação (informar se o participante FEZ algum teste para doenças infecto-contagiosas)	<input type="text"/>

Fonte: Autor

Temos então, a categoria de “Amostra biológica” (apêndice H), onde ficam registradas todas as informações de coleta da amostra, de extração e das alíquotas realizadas. Além de identificar onde as mesmas se encontram no espaço físico do laboratório.

5.3.2 Sessões específicas de prontuário

Nas sessões específicas, temos os questionamentos de prontuários específicos para Artrite (apêndice I), Câncer de mama (apêndice J), Lúpus (apêndice K) e Psoríase (apêndice L). Todos possuem em comum um campo de histórico obtido através de prontuário e um histórico que conta com a narrativa do participante. Este último se faz importante, pois ao trabalhar com público e com o agravante de se fazer um questionário completo de doenças, que muitas vezes fara com que a pessoa se lembre de coisas e momentos dolorosos, e necessário demonstrar empatia e abrir acesso para que o indivíduo possa se expressar da forma como desejar.

6. CONSIDERAÇÕES FINAIS

Computadores cada vez mais rápidos e mais baratos permitem abordar problemas, literalmente, inimagináveis há poucos anos. Os métodos e a dimensão dos problemas abordados por um aluno de iniciação científica serão, em sua maioria, totalmente obsoletos ao final de seu doutorado. A cada ano que passa problemas cada vez mais complexos podem ser abordados, de forma mais completa, e mais pesquisadores com menos recursos financeiros podem trabalhar nestas áreas de pesquisa, o que torna a bioinformática uma das áreas do conhecimento mais acessíveis para pesquisadores em início de carreira (VERLI, 2014).

Outro aspecto a ser levado em conta no processo de levantamento de dados e constituições de amostras é a fidedignidade. Este conceito está relacionado com a qualidade e precisão dos dados. Ao usar questionários para levantamento de dados, ficamos expostos a erros decorrentes de informações com pouca precisão e erros no preenchimento do formulário por parte do pesquisador (ARANGO, 2011). Os questionários construídos para o LAPOGEDb minimizam as chances de erros de preenchimento e auxiliam o participante a responder de forma mais assertiva possível.

Perspectivas:

- ✓ Alimentação do banco de dados de forma padronizada, através da distribuição de tabelas normalizadas para preenchimento dos dados obtidos através das pesquisas realizadas no LAPOGE. Desta forma será possível otimizar e organizar todas as informações que já foram obtidas e, através disso, verificar quais marcadores ainda não foram testados para cada amostra.
- ✓ A organização sistematizada das amostras, com etiquetas térmicas para armazenagem no estoque.
- ✓ Implementação de questionários padronizados e disponíveis no banco para preenchimento, o que tornará possível o acesso ao mesmo através de *tablets*, *smartphones* ou acesso remoto *online* em computadores portáteis onde serão preenchidos e incorporados automaticamente ao LAPOGEDb. O uso desta ferramenta, substituiria a necessidade de realizar questionários físicos para a coleta dos dados, otimizando o tempo, ocupação de espaço físico e reduzindo a margem de erros.
- ✓ Elaboração de sistema de busca (query) utilizando combinações.

REFERÊNCIAS

ALVES, S. M. REBES REVISTA BRASILEIRA DE EDUCAÇÃO E SAÚDE GVAA -GRUPO VERDE DE AGROECOLOGIA E ABELHAS -POMBAL -PB Artigo de Revisão A bioinformática e sua importância para a biologia molecular *Bioinformatics and its importance for molecular biology*. n. 4, p. 18–25, 2013.

ARANGO, Héctor G. **Bioestatística**: teórica e computacional. 3. ed. Rio de Janeiro: Guanabara Koogan, 2011. 438 p. ISBN 978-85-277-1558-4

ATTWOOD, T. K. G. A E. N.-E. et al. Concepts, Historical Milestones and the Central Place of Bioinformatics in Modern Biology: A European Perspective. **Bioinformatics - Trends and Methodologies**, p. 1–36, 2011.

AUTON, A. et al. A global reference for human genetic variation. **Nature**, v. 526, n. 7571, p. 68–74, 2015.

BAXEVANIS, A. D. The importance of biological databases in biological discovery. **Current Protocols in Bioinformatics**, v. Chapter 1, p. Unit 1.1, 2011.

CLARKE, L. et al. The 1000 Genomes Project: data management and community access. **Nature Methods**, v. 9, n. 5, p. 459–462, 2012.

CORRÊA, M. V. O admirável Projeto Genoma Humano. **Physis: Revista de Saúde Coletiva**, v. 12, n. 2, p. 277–299, 2002.

COSTA, R. C. DA. A EVOLUÇÃO E REVOLUÇÃO DA ESCRITA: UM ESTUDO COMPARATIVO. p. 1–99, 2002.

DATE, C. J. **An Introduction to Database Systems** (8th Edition). Addison-Wesley, 2003.

DEBORTOLI, G. **Elaboração de um painel de marcadores de ancestralidade para estudos caso-controle e organização das informações em banco de dados local**. Dissertação, UFSC, 2013.

DIAS, C. A. Hipertexto: evolução histórica e efeitos sociais. **Ciência da Informação**, v. 28, n. 3, p. 269–277, 1999.

ELMASRI, R.; NAVATHE, S.B. 2010. **Sistema de banco de dados**. 4^a ed. Pearson Education, São Paulo, Brasil, 724pp.

ELTABAKH, M. Y. **bdbms – A Database Management System for Biological Data**. 2006.

HAAG, C. A arca humana num diúvio de dados. 2011.

HIGGINS, D.; TAYLOR, W. Bioinformatics: Sequence, Structure and Databanks. **Briefings in Bioinformatics**, v. 2, n. 2, p. 202–204, 2001.

JORDAN, D. M.; RAMENSKY, V. E.; SUNYAEV, S. R. Human allelic variation: perspective from protein function, structure, and evolution. **Current opinion in structural biology**, v. 20, n. 3, p. 342–50, 2010.

HILBERT, M.; LÓPEZ, P. The World ' s Technological Capacity. **Sciencemag.org**, v. 332, n. April, p. 60–65, 2011.

JOSÉ, C.; RIBEIRO, S. Big Data: os novos desafios para o profissional da informação. v. 1, n. 1, p. 96–105, 2014.

MOREIRA, L. M. **Ciencias Genômicas**. Ribeirão Preto: [s.n.]. v. 53, 2015.

.ROBBINS, R. J. **Biological Databases: A New Scientific Literature**. 1994.

SIMPSON, A. J., F. C. REINACH, et al. (2000). “The genome sequence of the plant pathogen *Xylella fastidiosa*. The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing and Analysis.” **Nature** 406(6792): 151-159.

SILVA, M. S. **Construindo Sites com CSS e (X)HTML**. Novatec, 2008.

STEIN, L. D. Integrating biological databases. **Nature reviews. Genetics**, v. 4, n. 5, p. 337–45, 2003.

STEPHENS, Z. D. et al. Big data: Astronomical or genetical? **PLoS Biology**, v. 13, n. 7, p. 1–11, 2015.

TAHAGHOGHI, S.; WILLIAMS, H. **Learning MySQL**. O'Reilly, 2007

TURING, A. M. On computable numbers, with an application to the entscheidungsproblem. a correction. **Proceedings of the London**

Mathematical Society, v. s2-43, n. 1, p. 544–546, 1938.

VERLI, H. **Bioinformática da Biologia à flexibilidade molecular**. Porto Alegre: [s.n.].

VIA, M. et al. The 1000 Genomes Project: New opportunities for research and social challenges. **Genome Medicine**, v. 2, n. 1, p. 3, 2010.

WILTON, P.; COLBY, J. **Beginning SQL**. Wrox, 2005.

APÊNDICE A – Trabalhos realizados no LAPOGE, organizados por nome do autor

Ano	Tipo	Autor	Título do trabalho
2009	Dissertação	Aline Fernanda Rodrigues Sereia	Identificação de polimorfismos em genes de reparo de dna e de detoxificação como possíveis marcadores de susceptibilidade ao câncer de mama
2012	TCC	Amanda Firmino	Estudo de associação da alu ygeind114pbdo gene hla-gem pacientes com psoríase do estado de santa catarina
2004	TCC	Ana Cristina Jorck	Polimorfismos dos genes cyp1a1 e gstm1 em uma amostra de mulheres da população do estado de santa catarina
2004	TCC	Ana Frederica Sutter Latorre	Polimorfismo do microssatélite do loco lpl, nas populações semi-isoladas de valongo (porto belo) e costa da lagoa (florianópolis), em santa catarina. 2004.
2005	TCC	Ana Márcia Teixeira Moraes	Avaliação do gene tp53em câncer de mama
2015	TCC	Andressa de Lima	Ascendência genética matrilinear da grande florianópolis
2012	Dissertação	Bibiana Sgorla de Almeida	Polimorfismos do gene hla-gem pacientes com psoríase
2012	TCC	Bruna Viviane Vaz	Uso dos minists non-codis 01 na prática forense: estudo populacional de santa catarina

2013	TCC	Cileide Brasil da Silva	Polimorfismos do gene hla-g em pacientes com artrite reumatoide e controles no estado de santa catarina. 2013. Trabalho de conclusão de curso.
2008	TCC	Cíntia Callegari Coêlho	Análise de associação de dois polimorfismos do gene hfrapredisposição ao câncer de mama em mulheres de santa catarina
2011	Dissertação	Cíntia Callegari Coêlho	Associação de genes kir ao lúpus eritematoso sistêmico em santa catarina
2004	TCC	Clarissa Torresan	Estudo do polimorfismo do loco f13b em populações isoladas em santa catarina: valongo (porto belo) e costa da lagoa (florianópolis)
2015	TCC	Clisten Fatima Staffen	Identificação biológica e adulteração de pescados comercializados em peixarias de florianópolis/sc através de dna barcoding
2007	TCC	Dianne Galiotto	Câncer de mama em santa catarina: análise de polimorfismo do gene gsp1 em um estudo caso-controle
2011	Dissertação	Dianne Galiotto	Frequência de mutações no gene tp53 em tumores de mama em pacientes de santa catarina
2012	TCC	Emily Bruna Justino	Implementação de marcadores minisatélites para obtenção de frequências alélicas na população de santa catarina
1997	TCC	Fabiana Cristine da Rosa	Identificação de proteínas plasmáticas nas comunidades da costa da lagoa e de são joão do rio vermelho, ilha de santa catarina, sul do brasil

1996	TCC	Gabriela de Medeiros Saldanha	Identificação da variabilidade dos grupos sanguíneos eritrocitários nas comunidades da costa da lagoa e rio vermelho
2012	TCC	Gabriela Duarte Karasiak	Análise dos polimorfismos da região 3' não traduzida do gene hla-g e inserção alu em pacientes com lúpus eritematoso sistêmico na população do estado de Santa Catarina
2015	Dissertação	Gabriela Duarte Karasiak	Análise dos polimorfismos da região 3' não traduzida do gene hla-g e inserção alu em pacientes de Santa Catarina com histórico de artrite reumatoide
2013	Dissertação	Guilherme Debortoli	Elaboração de um painel de marcadores de ancestralidade para estudos caso-controle e organização das informações em banco de dados local
2013	TCC	Günther Gerhard Gerent	Utilização de dois marcadores informativos de ancestralidade para estudos de caso-controle
2013	TCC	Jimena Pereira Rodrigues Kirchner	Inserção alu em pacientes com artrite reumatoide
2014	TCC	Larissa da Silva Luzietti	Associação do polimorfismo c1858 do gene ptpn22 com lúpus eritematoso sistêmico em Santa Catarina, Brasil
2012	TCC	Leili Daiane Hausmann	Análise epidemiológica e de associação de polimorfismos do complexo principal de histocompatibilidade com o câncer de mama em mulheres de Santa Catarina

2015	Dissertação	Leili Daiane Hausmann	Câncer de mama: busca de mirnas que controlam expressão de hla-g, uma molécula imunorreguladora, como biomarcador para diagnósticos e opções terapêuticas. Início: 2015
1999	TCC	Leonardo Alves Júnior	
2004	TCC	Lia kubelka Fernandes de Carlos Back	Marcadores de susceptibilidade e auxílio diagnóstico no lupus eritematoso sistêmico
2007	Dissertação	Lia kubelka Fernandes de Carlos Back	Lupus eritematoso sistêmico: pesquisa de marcadores moleculares de susceptibilidade e prognóstico
2014	Tese	Lia kubelka Fernandes de Carlos Back	Associação de polimorfismos em genes envolvidos na regulação da atividade das células nk (natural killer) no desenvolvimento do câncer de mama
2004	TCC	Lílian Pires Barbata	Estudo do polimorfismo dos locos hla-dra, hla-drb1, hla-drb3, hla-drb5, hla-dqa1 e hla-dqb1 na população isolada de valongo, no estado de santa catarina
2009	TCC	Luisa Matos do Canto	Estudo epidemiológico e de associação do polimorfismo do gene pdcd1 à artrite reumatóide e ao lúpus eritematoso sistêmico em santa catarina
2015	TCC	Mari Dalva Staffen	Uso da técnica molecular dna barcodena identificação de substituições fraudulentas de pescados comercializados em restaurantes de culinária japonesa de florianópolis/ sc

2004	TCC	Maria Fernanda Valverde da Silva	Polimorfismo dos locos hla-a, -b e -drb1 em doadores de medula óssea no estado de santa catarina
2012	TCC	Maria Luiza Guimarães de Oliveira	Análise do polimorfismo ptpn22c1858tpor high resolution meltingem pacientes comartrite reumatoide
2012	Dissertação	Mariáh Damiani da Silva	Polimorfismos da região 3' não traduzida do gene hla-gem pacientes com câncer de mama e controles no estado de santa catarina
2005	TCC	Mariana Chagas/ Ana Paula Rigon	Análise dos polimorfismos dos genes cyp1a1,gstm1e gstt1em mulheres sadias e com câncer de mama em santa catarina
2015	TCC	Mariana Londero Becker	Estudo genético da 3'utr do gene mhc-gem primatas não humanos da família cebidae a partir de metodologias utilizadas em primatas humanos
2009	TCC	Mayara Delagnelo Medeiros	Polimorfismo do gene gstp1 em um estudo epidemiológico e de associação com lúpus eritematoso sistêmico e artrite reumatóide em santa catarina
2011	TCC	Patrícia Amorim da Cunha	Influência depolimorfismoem genes relacinados à obesidade -ftoc mc4r-sobreo câncer de mama em mulheres do estado de santa catarina
2014	Tese	Sandra Regina Rachadel Torres	Avaliação da estrutura genética da população atual de santa catarina com diferentes marcadores moleculares para aplicação na genética forense

2015	TCC	Saulo José dos Reis	Análises filogenéticas da região mhc em primatas não-humanos. Revisão bibliográfica da evolução do gene mhc-g
2007	TCC	Tiago Moreti	Polimorfismos do gene cyp11a1 e sua associação com mulheres portadoras de câncer de mama em Santa Catarina
2009	Dissertação	Tiago Moreti	Identificação humana: uma proposta metodológica para obtenção de DNA de ossos e implementação de banco de dados de frequências alélicas de STRs autossômicos na população de Santa Catarina
2009	TCC	Ticiano Della Justina Farias	Associação do gene IL-18 com artrite reumatóide em um estudo epidemiológico no estado de Santa Catarina: estudo caso-controle
2012	Dissertação	Ticiano Della Justina Farias	Análise epidemiológica e de associação dos genes KIR com artrite reumatóide
1999	TCC	Yara Costa Neto Muniz	Novos dados referentes às comunidades isoladas da Costa da Lagoa e de São João do Rio Vermelho, na Ilha de SC
2003	Dissertação	Yara Costa Neto Muniz	Comunidades semi-isoladas fundadas por açorianos na Ilha de Santa Catarina
2008	Tese	Yara Costa Neto Muniz	Marcadores genéticos em duas comunidades semi-isoladas fundadas por açorianos na Ilha de Santa Catarina

APÊNDICE B - Lista de marcadores utilizados nos trabalhos realizados no LAPOGE ao decorrer dos 22 anos de sua existência, organizados por nome do marcador

Nome	Tipo	Referência
A1298C (MTHFR)	SNPs	Coelho, 2008
A4889G	SNPs	Back, 2007.
AluyHG	InDel	Kirchner, 2013; Firmino, 2012; Karasiak, 2012; Hausmann, 2012
APO	InDel	Debortoli, 2013; Gerent, 2013
Arg194Trp	SNPs	Sereia, 2009.
Arg399Gln	SNPs	Sereia, 2009.
AT3-I/D	InDel	Debortoli, 2013.
C1858T (PTPN22)	SNPS	Luzietti, 2014; Oliveira, 2013
C4887A	SNPs	Back, 2007.
C677T (MTHFR)	SNPs	Coelho, 2008
CSF1PO	STR	Moreti, 2009.
CYP1A1	SNPs	Moraes, 2005; Rigon & Chagas, 2005; Jorck, 2004; Back 2004.
D10S1248	miniSTRs	Torres, 2014; Vaz, 2012
D10S2325	STRs	Torres, 2014.
D12S391	STRs	Torres, 2014.
D12S391	miniSTRs	Torres, 2014.
D13S317	STR	Moreti, 2009.
D14S1434	miniSTRs	Vaz, 2012.
D16S539	STR	Moreti, 2009.

D18S51	STRs	Torres, 2014; Moreti 2009.
D19S433	STR	Moreti, 2009.
D1S1656	miniSTRs	Torres, 2014.
D1S1677	miniSTRs	Justino, 2013
D21S11	STR	Moreti, 2009.
D21S2055	STRs	Torres, 2014.
D22S1045	miniSTRs	Torres, 2014; Vaz, 2012
D2S1338	STR	Moreti, 2009.
D2S1360	STRs	Torres, 2014.
D2S441	miniSTRs	Torres, 2014; Justino, 2012
D3S1358	STR	Moreti, 2009.
D3S1744	STRs	Torres, 2014.
D4S2364	miniSTRs	Justino, 2012
D4S2366	STRs	Torres, 2014.
D5S2500	STRs	Torres, 2014.
D5S818	STR	Moreti, 2009.
D6S474	STRs	Torres, 2014.
D7S1517	STRs	Torres, 2014.
D7S820	STR	Moreti, 2009.
D8S1132	STRs	Torres, 2014.
D8S1179	STR	Moreti, 2009.
DXS10074	X-STRs	Torres, 2014.
DXS10079	X-STRs	Torres, 2014.

DXS10101	X-STRs	Torres, 2014.
DXS10103	X-STRs	Torres, 2014.
DXS10134	X-STRs	Torres, 2014.
DXS10135	X-STRs	Torres, 2014.
DXS10146	X-STRs	Torres, 2014.
DXS10148	X-STRs	Torres, 2014.
DXS7132	X-STRs	Torres, 2014.
DXS7423	X-STRs	Torres, 2014.
DXS8378	X-STRs	Torres, 2014.
DYS19	Y-STRs	Torres, 2014.
DYS385a/b	Y-STRs	Torres, 2014.
DYS389I/II	Y-STRs	Torres, 2014.
DYS390	Y-STRs	Torres, 2014.
DYS391	Y-STRs	Torres, 2014.
DYS392	Y-STRs	Torres, 2014.
DYS393	Y-STRs	Torres, 2014.
DYS437	Y-STRs	Torres, 2014.
DYS438	Y-STRs	Torres, 2014.
DYS439	Y-STRs	Torres, 2014.
DYS448	Y-STRs	Torres, 2014.
DYS456	Y-STRs	Torres, 2014.
DYS458	Y-STRs	Torres, 2014.
DYS635 (GATA C4)	Y-STRs	Torres, 2014.

F13B	SNPs	Torresan, 2004.
FGA	STR	Moreti, 2009.
FTO	SNPs	da Cunha, 2011
FTO	SNPs	da Cunha, 2011
GATA H4	Y-STRs	Torres, 2014.
GSTM1	SNPs	Back, 2007; Moraes, 2005; Rigon & Chagas, 2005; Jorck, 2004.
GSTP1 A313G	SNPs	Medeiros, 2009; Galiotto, 2007.
GSTT1	SNPs	Back, 2007; Moraes, 2005; Rigon & Chagas, 2005.
HLA-A	SNPs	da Silva, 2004.
HLA-B	SNPs	da Silva, 2004.
HLA-DQA1	SNPs	Barbeta, 2004.
HLA-DQB1	SNPs	Barbeta, 2004.
HLA-DRA	SNPs	Barbeta, 2004.
HLA-DRB1	SNPs	da Silva, 2004; Barbeta, 2004.
HLA-DRB3	SNPs	Barbeta, 2004.
HLA-DRB5	SNPs	Barbeta, 2004.
HLA-G +3001 T/G	SNPs	Karasiak, 2015.
HLA-G +3003 T/C	SNPs	Karasiak, 2015; Almeida, 2012;
HLA-G +3010 C/G	SNPs	Karasiak, 2015; Almeida, 2012;
HLA-G +3027 A/C	SNPs	Karasiak, 2015; Almeida, 2012;
HLA-G +3035 C/T	SNPs	Karasiak, 2015; Almeida, 2012;
HLA-G +3142 G/C	SNPs	Karasiak, 2015; Almeida, 2012; Damiani, 2012

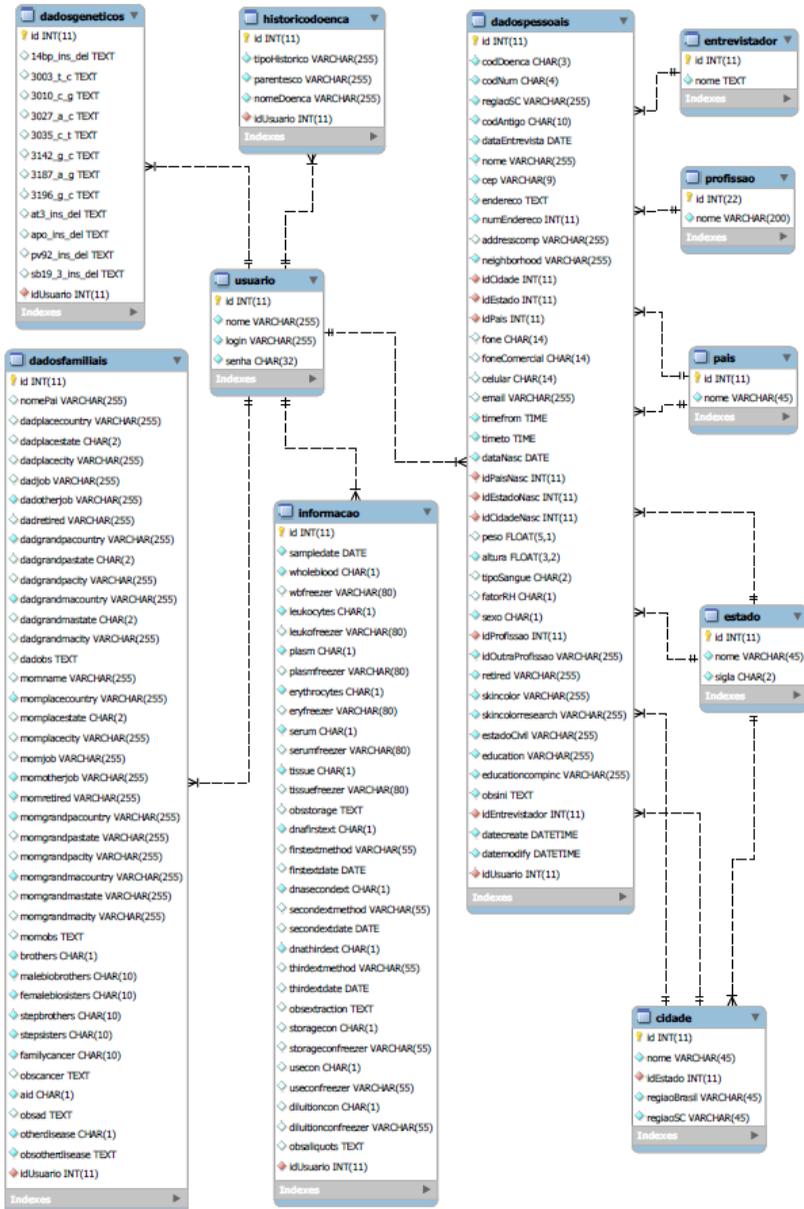
HLA-G +3187 A/G	SNPs	Karasiak, 2015; Almeida, 2012.
HLA-G +3196 G/C	SNPs	Karasiak, 2015; Almeida, 2012.
HLA-G +3227 A/G	SNPs	Karasiak, 2015.
HLD101	InDel	Torres, 2014.
HLD111	InDel	Torres, 2014.
HLD114	InDel	Torres, 2014.
HLD118	InDel	Torres, 2014.
HLD122	InDel	Torres, 2014.
HLD124	InDel	Torres, 2014.
HLD125	InDel	Torres, 2014.
HLD128	InDel	Torres, 2014.
HLD131	InDel	Torres, 2014.
HLD133	InDel	Torres, 2014.
HLD136	InDel	Torres, 2014.
HLD39	InDel	Torres, 2014.
HLD40	InDel	Torres, 2014.
HLD45	InDel	Torres, 2014.
HLD48	InDel	Torres, 2014.
HLD56	InDel	Torres, 2014.
HLD58	InDel	Torres, 2014.
HLD6	InDel	Torres, 2014.
HLD64	InDel	Torres, 2014.
HLD67	InDel	Torres, 2014.

HLD70	InDel	Torres, 2014.
HLD77	InDel	Torres, 2014.
HLD81	InDel	Torres, 2014.
HLD83	InDel	Torres, 2014.
HLD84	InDel	Torres, 2014.
HLD88	InDel	Torres, 2014.
HLD92	InDel	Torres, 2014.
HLD93	InDel	Torres, 2014.
HLD97	InDel	Torres, 2014.
HLD99	InDel	Torres, 2014.
HPRTB	X-STRs	Torres, 2014.
HV1	Sequenciamento	Torres, 2014.
HVS-1	Sequenciamento	de Lima, 2015
IL18-137 (G/C)	SNPs	Back, 2014; Farias, 2009; Back 2007
IL18-607 (C/A)	SNPs	Back, 2014; Farias, 2009; Back 2007
InDel14pb	InDel	Karasiak, 2015; Firmino, 2012; Almeida 2012; Karasiak, 2012; Hausmann, 2012; Damiani, 2012
KIR2DL1	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011.
KIR2DL2	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011.
KIR2DL3	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011.
KIR2DL4	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011.

KIR2DL5	Pres/Aus	Back, 2014; Farias, 2012
KIR2DP1	Pres/Aus	Farias, 2012; Coelho, 2011
KIR2DS1	Pres/Aus	Farias, 2012; Coelho, 2011
KIR2DS2	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011
KIR2DS3	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011
KIR2DS4	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011
KIR2DS5	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011
KIR3DL1	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011
KIR3DL2	Pres/Aus	Back, 2014; Farias, 2012
KIR3DL3	Pres/Aus	Back, 2014; Farias, 2012
KIR3DP1	Pres/Aus	Back, 2014; Farias, 2012
KIR3DS1	Pres/Aus	Back, 2014; Farias, 2012; Coelho, 2011
MC4R	SNPs	da Cunha, 2011
NKG2D	SNPs	Back, 2014.
PD1.3 ou +7146 PTPN22	SNPs	do Canto, 2009; Back 2007.
PV92	InDel	Debortoli, 2013; Gerent, 2013
Sb19.3	InDel	Debortoli, 2013.
SE33	STRs	Torres, 2014.
T6235C	SNPs	Back, 2007.

TH01	miniSTRs	Torres, 2014; Moreti 2009.
TP53 éxon 5	Sequenciamento	Galiotto, 2011; Moraes 2005.
TP53 éxon 6	Sequenciamento	Galiotto, 2011; Moraes 2005.
TP53 éxon 7	Sequenciamento	Galiotto, 2011; Moraes 2005.
TP53 éxon 8	Sequenciamento	Galiotto, 2011; Moraes 2005.
TPOX	STR	Moreti, 2009.
Val158Met	SNPs	Sereia, 2009.
Val-9Ala	SNPs	Sereia, 2009.
vWA	STR	Moreti, 2009.

APÊNDICE C- Estrutura relacional do LAPOGEDb



APÊNDICE D - Campo de dados familiares do LAPOGEDb.

IDENTIFICAÇÃO INICIAL	
CÓDIGO	<?php echo \$_SESSION["diseasepopcod"]; \$_SESSION["questcode"]; ?>

INFORMAÇÕES PATERNAS	
Possui informações paternas?	<input type="radio"/> Não <input type="radio"/> Sim
Nome do pai do entrevistado	<input type="text"/> manter em sigilo
Cidade/Estado de nascimento	-- seleçione -- -- Primeiro seleçione o estado --
Profissão	-- seleçione -- <input type="checkbox"/> Aposentado <input type="checkbox"/> Falecido
Onde nasceu o avô paterno do probando	<input type="checkbox"/> Não sabe informar -- Primeiro seleçione o estado -- selected="selected" -- seleçione --
Onde nasceu a avó paterna do probando	<input type="checkbox"/> Não sabe informar -- Primeiro seleçione o estado -- selected="selected" -- seleçione --
Observação	<input type="text"/>

INFORMAÇÕES MATERNAS	
Possui informações maternas?	<input type="radio"/> Não <input type="radio"/> Sim
Nome da mãe do entrevistado	<input type="text"/> manter em sigilo
Cidade/Estado de nascimento	-- seleçione -- -- Primeiro seleçione o estado --
Profissão	-- seleçione -- <input type="checkbox"/> Aposentada <input type="checkbox"/> Falecida
Onde nasceu o avô materno do probando	<input type="checkbox"/> Não sabe informar -- Primeiro seleçione o estado -- selected="selected" -- seleçione --
Onde nasceu a avó materna do probando	<input type="checkbox"/> Não sabe informar -- Primeiro seleçione o estado -- selected="selected" -- seleçione --
Observação	<input type="text"/>

INFORMAÇÕES FRATERNAIS	
Possui irmãos?	<input type="radio"/> Não <input type="radio"/> Sim
Biológicos	<input type="checkbox"/> Não possui irmãos <input type="checkbox"/> Não possui irmãs
Adotivos	<input type="checkbox"/> Não possui irmãos <input type="checkbox"/> Não possui irmãs
Observação	<input type="text"/>

ANTECEDENTES FAMILIARES	
Histórico de Câncer familiar?	<input type="radio"/> Não <input type="radio"/> Sim
Descrição familiar	-- seleçione o parente -- -- seleçione o tipo -- + um parente >>
Observação	<input type="text"/>
Histórico doença autoimune?	<input type="radio"/> Não <input type="radio"/> Sim
Descrição familiar	-- seleçione o parente -- -- seleçione o tipo -- + um parente >>
Observação	<input type="text"/>
Histórico de outras doenças?	<input type="radio"/> Não <input type="radio"/> Sim
Descrição familiar	-- seleçione o parente -- -- seleçione o tipo -- + um parente >>
Observação	<input type="text"/>

[Limpar Tudo](#)
[Salvar / Próxima Etapa](#)

APÊNDICE E - Campo de dados epidemiológicos do LAPOGEDb

HÁBITOS TABAGISTAS

Você fuma?	<input type="radio"/> Não, nunca fumei <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, fumo atualmente
Que tipo de cigarro?	<input type="checkbox"/> Cigarro comercial <input type="checkbox"/> Cigarro caseiro <input type="checkbox"/> Charuto <input type="checkbox"/> Cachimbo <input type="checkbox"/> Outros
Qual a quantidade?	<input type="checkbox"/> Um cigarro <input type="checkbox"/> 2 a 3 cigarros <input type="checkbox"/> 3 a 5 cigarros <input type="checkbox"/> mais que 5 cigarros
Com que frequência?	<input type="radio"/> Esporadicamente <input type="radio"/> Uma vez por dia <input type="radio"/> 2 a 5 vezes por dia <input type="radio"/> 5 a 10 vezes por dia <input type="radio"/> Mais que 10 vezes por dia
Convive com alguém que fuma?	<input type="radio"/> Não <input type="radio"/> Sim
Como você classifica o hábito tabagista desta pessoa?	<input type="radio"/> leve <input type="radio"/> moderado <input type="radio"/> acentuado <input type="radio"/> pesado
Observação	

HÁBITOS ALCOÓLICOS

Você ingere bebidas alcoólicas?	<input type="radio"/> Não, nunca bebi <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, bebo atualmente
Que tipo de bebida?	<input type="checkbox"/> Cerveja <input type="checkbox"/> Vinho <input type="checkbox"/> Cachaça <input type="checkbox"/> Outros
Qual a quantidade?	<input type="checkbox"/> no máximo 1 copo <input type="checkbox"/> entre 2 e 5 copos <input type="checkbox"/> entre 5 e 9 copos <input type="checkbox"/> mais que 10 copos
Com que frequência?	<input type="radio"/> Esporadicamente <input type="radio"/> Nos fins de semana <input type="radio"/> Uma vez por dia <input type="radio"/> 2 a 3 vezes por dia <input type="radio"/> 3 a 5 vezes por dia <input type="radio"/> Mais que 5 vezes por dia
Convive com alguém que bebe?	<input type="radio"/> Não <input type="radio"/> Sim
Como você classifica o hábito tabagista desta pessoa?	<input type="radio"/> leve <input type="radio"/> moderado <input type="radio"/> acentuado <input type="radio"/> pesado
Observação	

HÁBITOS DE ADIÇÃO DE DROGAS (não considerar tabagismo, alcoolismo ou uso de medicamentos)

Você consome alguma droga?	<input type="radio"/> Não, nunca usei <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, uso atualmente
Que tipo de droga?	<input type="checkbox"/> Maconha <input type="checkbox"/> Cocaína <input type="checkbox"/> Drogas sintéticas <input type="checkbox"/> Chás e/ou cogumelos <input type="checkbox"/> Outros
Com que frequência?	<input type="radio"/> Esporadicamente <input type="radio"/> Nos fins de semana <input type="radio"/> Uma vez por dia <input type="radio"/> 2 a 3 vezes por dia <input type="radio"/> 3 a 5 vezes por dia <input type="radio"/> Mais que 5 vezes por dia
Convive com alguém que consome?	<input type="radio"/> Não <input type="radio"/> Sim
Como você classifica o hábito de adição desta pessoa?	<input type="radio"/> leve <input type="radio"/> moderado <input type="radio"/> acentuado <input type="radio"/> pesado
Observação	<div style="border: 1px solid black; height: 30px;"></div>

PRÁTICA DE EXERCÍCIOS FÍSICOS

Você pratica exercícios físicos regularmente?	<input type="radio"/> Não, nunca pratiquei <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, pratico atualmente
Que tipo de atividade?	<input type="checkbox"/> Academia (musculação) <input type="checkbox"/> Caminhada/ Esteira <input type="checkbox"/> Natação/ Hidroginástica <input type="checkbox"/> Corrida <input type="checkbox"/> Bicicleta/ bicicleta ergométrica <input type="checkbox"/> Jogos coletivos <input type="checkbox"/> Artes marciais <input type="checkbox"/> Outros
Com que frequência?	<input type="radio"/> Esporadicamente <input type="radio"/> Nos fins de semana <input type="radio"/> Uma vez por semana <input type="radio"/> 2 a 3 vezes por semana <input type="radio"/> 3 a 5 vezes por semana <input type="radio"/> Todos os dias da semana
Qual a duração da atividade?	<input type="radio"/> menos de meia hora <input type="radio"/> menos de uma hora <input type="radio"/> uma hora <input type="radio"/> mais de uma hora
Observação	<div style="border: 1px solid black; height: 30px;"></div>

APÊNDICE F – Campo de dados biológicos do LAPODEdb

ALERGIAS

Alimentos	<input type="radio"/> Não <input type="radio"/> Sim	
Medicamentos	<input type="radio"/> Não <input type="radio"/> Sim	
Outras	<input type="radio"/> Não <input type="radio"/> Sim	

HISTÓRICO HORMONAL

Menarca	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não <input type="radio"/> Sim	com _____ anos
Menopausa	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não <input type="radio"/> Sim	com _____ anos
Histerectomia	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não <input type="radio"/> Sim	com _____ anos
Ovarietomia	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não <input type="radio"/> Sim	com _____ anos
Reposição Hormonal	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não, nunca fiz <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, estou fazendo	_____ por quanto tempo? _____ há quanto tempo parou?
Outros hormônios	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Não, nunca usei <input type="radio"/> Sim, mas parei <input type="radio"/> Sim, estou usando	_____ por quanto tempo? _____ qual motivo? _____ há quanto tempo parou?
Observação		

USO DE CONTRACEPTIVOS

Utiliza anticoncepcional?	<input type="radio"/> Não se aplica por ser homem <input type="radio"/> Nunca usei <input type="radio"/> Usei mas parei <input type="radio"/> Sim, estou usando
Que tipo?	<input type="checkbox"/> Oral <input type="checkbox"/> Injetável <input type="checkbox"/> Adesivo <input type="checkbox"/> Outros
Informações complementares	_____ Há quanto tempo usa? _____ se parou, por quanto tempo usou? _____ se parou, qual o motivo? _____ se parou, há quanto tempo parou?
Observações	

PARIDADE		
Multiparidade	<input type="checkbox"/> Não se aplica por ser homem <input type="checkbox"/> Sim, por opção <input type="checkbox"/> Sim, por limitações biológicas <input type="checkbox"/> Não é nulípara	
Uniparidade	<input type="checkbox"/> Não se aplica por ser homem <input type="checkbox"/> Sim, por opção <input type="checkbox"/> Sim, por limitações biológicas <input type="checkbox"/> Não é unípara	Gestação com _____ anos
Multiparidade	<input type="checkbox"/> Não se aplica por ser homem <input type="checkbox"/> Sim <input type="checkbox"/> Não é multipara	<input type="text"/> Número de gestações <input type="text"/> Filhos (masculino) <input type="text"/> Filhas (feminino) <input type="text"/> anos
Observações Adicionais	<input type="text"/>	

HISTÓRICO DE ABORTOS	
Você tem histórico de aborto?	<input type="checkbox"/> Não se aplica por ser homem <input type="checkbox"/> Não tenho conhecimento de ter abortado <input type="checkbox"/> Sim, já abortei
Que tipo?	<input type="checkbox"/> Provocado <input type="checkbox"/> Acidental
Observações	<input type="text"/>

HISTÓRICO DE AMAMENTAÇÃO	
Você amamentou?	<input type="checkbox"/> Não se aplica por ser homem <input type="checkbox"/> Não amamentei pois não tive filhos <input type="checkbox"/> Não amamentei meus filhos por decisão própria <input type="checkbox"/> Não amamentei meus filhos por causa do trabalho <input type="checkbox"/> Não amamentei meus filhos por limitações biológicas
Quanto tempo?	<input type="text"/>
Observações	<input type="text"/>

APÊNDICE G – Campo de dados médicos do LAPOGEDb

HISTÓRICO CARDIOLÓGICO

Algum relato cardiológico?	<input type="checkbox"/> Angina <input type="checkbox"/> Cateterismo prévio <input type="checkbox"/> Doença Coronária <input type="checkbox"/> Hipertensão Arterial Sistêmica <input type="checkbox"/> Infarto Agudo do Miocárdio <input type="checkbox"/> Revascularização do miocárdio <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

HISTÓRICO AUTOIMUNE

Algum relato autoimune?	<input type="checkbox"/> Artrite reumatoide <input type="checkbox"/> Diabetes <input type="checkbox"/> Doença de Crohn <input type="checkbox"/> Lupus Eritematoso Sistêmico <input type="checkbox"/> Pênfigo <input type="checkbox"/> Psoríase <input type="checkbox"/> Vitiligo <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

HISTÓRICO NEUROLÓGICO

Algum relato neurológico?	<input type="checkbox"/> Ataque Isquêmico Transitório <input type="checkbox"/> Acidente Vascular Cerebral <input type="checkbox"/> Ateroma em carótida <input type="checkbox"/> Mal de Alzheimer <input type="checkbox"/> Mal de Parkinson <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

HISTÓRICO PSIQUICO

Tem depressão?	<input type="radio"/> Não <input type="radio"/> Sim	<input type="text"/> anos
Teve depressão?	<input type="radio"/> Não <input type="radio"/> Sim	com <input type="text"/> anos
Utiliza medicamentos?	<input type="radio"/> Não <input type="radio"/> Sim	<input type="text"/> Quais <input type="text"/> Que dose? <input type="text"/> Há quanto tempo?
Utilizou medicamentos?	<input type="radio"/> Não <input type="radio"/> Sim	<input type="text"/> Quais <input type="text"/> Que dose? <input type="text"/> Por quanto tempo?
Tem ansiedade?	<input type="radio"/> Não <input type="radio"/> Sim	<input type="text"/> anos
Teve ansiedade?	<input type="radio"/> Não <input type="radio"/> Sim	com <input type="text"/> anos

Utiliza medicamentos?	<input type="radio"/> Não <input type="radio"/> Sim	<input type="text"/> <input type="text"/> <input type="text"/>	Quais Que dose? Há quanto tempo?
Utilizou medicamentos	<input type="radio"/> Não <input type="radio"/> Sim	<input type="text"/> <input type="text"/> <input type="text"/>	Quais Que dose? Por quanto tempo?
Observação	<input type="text"/>		

HISTÓRICO ONCOLÓGICO

Algum relato oncológico?	<input type="checkbox"/> Colo de útero <input type="checkbox"/> Cólono reto <input type="checkbox"/> Esôfago <input type="checkbox"/> Fígado <input type="checkbox"/> Leucemia <input type="checkbox"/> Linfoma <input type="checkbox"/> Ovários <input type="checkbox"/> Pele - não melanoma <input type="checkbox"/> Pele - melanoma <input type="checkbox"/> Pâncreas <input type="checkbox"/> Próstata <input type="checkbox"/> Pulmões <input type="checkbox"/> Outro (descrever nas observações)	<input type="text"/> <input type="text"/> <input type="text"/>
Observação	<input type="text"/>	

HISTÓRICO DIVERSO

Algum relato neurológico?	<input type="checkbox"/> Asma <input type="checkbox"/> Dengue <input type="checkbox"/> Hipercolesterolemia <input type="checkbox"/> Hipertrigliceridemia <input type="checkbox"/> Hepatite <input type="checkbox"/> Aids <input type="checkbox"/> Osteoporose <input type="checkbox"/> Tuberculose <input type="checkbox"/> Outro (descrever nas observações)	<input type="text"/> <input type="text"/> <input type="text"/>
Informar todos os medicamentos que utiliza atualmente	<input type="text"/>	
Observação (informar se o participante fez algum teste para doenças infecto-contagiosas)	<input type="text"/>	

APÊNDICE H – Campo de amostra biológica do LAPOGEDb

IDENTIFICAÇÃO INICIAL

CÓDIGO

INFORMAÇÕES SOBRE A AMOSTRA BIOLÓGICA

Data da coleta	<input type="text"/>
Houve outra coleta?	<input type="radio"/> Não <input checked="" type="radio"/> Sim
Tipo de amostra coletada	<input type="checkbox"/> Sangue total <input type="checkbox"/> Saliva <input type="checkbox"/> Tecido <input type="checkbox"/> Tecido tumoral <input type="checkbox"/> Outros (especificar no campo de observações)
Tipo de amostra processada	<input type="text" value=">Geladeira"/> Sangue total <input type="text" value=">Geladeira"/> Leucócitos <input type="text" value=">Geladeira"/> Plasma <input type="text" value=">Geladeira"/> Soro <input type="text" value=">Geladeira"/> Tecido <input type="text" value=">Geladeira"/> Hemácias <input type="text" value=">Geladeira"/> Outros (especificar no campo de observações)

EXTRAÇÃO DE DNA

Primeira Extração	<input type="text" value=">Não se aplica"/>	Data da Extração	<input type="text" value="<?php echo \$_firstextdate; ? disabled='disabled'>"/>
Segunda Extração	<input type="text" value=">Não se aplica"/>	Data da Extração	<input type="text" value="<?php echo \$_secondextdate; ? disabled='disabled'>"/>
Terceira Extração	<input type="text" value=">Não se aplica"/>	Data da Extração	<input type="text" value="<?php echo \$_thirdextdate; ? disabled='disabled'>"/>
Observação	<input type="text"/>		

ALÍQUOTAS DE DNA

Concentração da amostra	<input checked="" type="checkbox"/> Concentração Estoque	<input type="text" value=">Geladeira"/>	<input type="text" value="valor quantificado"/>
Concentração da amostra	<input checked="" type="checkbox"/> Concentração de Uso	<input type="text" value=">Geladeira"/>	<input type="text" value="valor quantificado"/>
Concentração da amostra	<input checked="" type="checkbox"/> Diluição	<input type="text" value=">Geladeira"/>	<input type="text" value="valor quantificado"/>
Observação	<input type="text"/>		

APÊNDICE I – Campo de Artrite do do LAPOGEDb

HISTÓRICO DA DOENÇA	
Histórico (versão do participante)	<input type="text"/>
Histórico (do prontuário)	<input type="text"/>
Número do prontuário	<input type="text"/> manter em sigilo
Fator desencadeante	<input type="text"/>
DA3 28	<input type="text"/>
Fator reumatoide	<input type="text"/>
Classificação na primeira consulta	<input type="radio"/> Ausente <input type="radio"/> Leve <input type="radio"/> Moderada <input type="radio"/> Grave <input type="radio"/> Muito grave
Data do diagnóstico	<input type="text"/>

MANIFESTAÇÕES INICIAIS	
Manifestações iniciais	<input type="checkbox"/> Derrame articular <input type="checkbox"/> Dor articular <input type="checkbox"/> Febre <input type="checkbox"/> Rigidez matinal <input type="checkbox"/> Outro (descrever nas observações)
Articulações acometidas	<input type="checkbox"/> Ombro <input type="checkbox"/> Cotovelo <input type="checkbox"/> Punho <input type="checkbox"/> MCF <input type="checkbox"/> IFPM <input type="checkbox"/> MTF <input type="checkbox"/> Quadril <input type="checkbox"/> Joelho <input type="checkbox"/> Tornozelo <input type="checkbox"/> IFPP
Total de articulações envolvidas	- conte quantas articulações foram afetadas ▾
Manifestações Extrarticulares	<input type="checkbox"/> Pleurite <input type="checkbox"/> Vasculite reumatoide <input type="checkbox"/> Acometimento ocular <input type="checkbox"/> Acometimento renal <input type="checkbox"/> Sjögren secundário <input type="checkbox"/> Pericardite <input type="checkbox"/> Nódulos reumatóides <input type="checkbox"/> Acometimento pulmonar <input type="checkbox"/> Amiloidose <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

MANIFESTAÇÕES RECENTES

Manifestações recentes	<input type="checkbox"/> Derrame articular <input type="checkbox"/> Dor articular <input type="checkbox"/> Febre <input type="checkbox"/> Rigidez matinal <input type="checkbox"/> Outro (descrever nas observações)
Articulações acometidas	<input type="checkbox"/> Ombro <input type="checkbox"/> Cotovelo <input type="checkbox"/> Punho <input type="checkbox"/> MCF <input type="checkbox"/> IFPM <input type="checkbox"/> MTF <input type="checkbox"/> Quadril <input type="checkbox"/> Joelho <input type="checkbox"/> Tornozelo <input type="checkbox"/> IFPP
Total de articulações envolvidas	- conte quantas articulações foram afetadas ▼
Manifestações Extrarticulares	<input type="checkbox"/> Pleurite <input type="checkbox"/> Vasculite reumatoide <input type="checkbox"/> Acometimento ocular <input type="checkbox"/> Acometimento renal <input type="checkbox"/> Sjögren secundário <input type="checkbox"/> Pericardite <input type="checkbox"/> Nódulos reumatóides <input type="checkbox"/> Acometimento pulmonar <input type="checkbox"/> Amiloidose <input type="checkbox"/> Outro (descrever nas observações)
Observação	

TRATAMENTO

Corticosteroídes	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Metotrexato	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Sulfasalazena	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Antimélico	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Ciclofosfamida	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Infliximab	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Etanercept	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Aíne	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Analgésicos	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Outros	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência Não se aplica ▼
Observação					

HAQ - HEALTH ASSESSMENT QUESTIONNAIRE (grau de dificuldade)

Você é capaz de?	com quanta dificuldade?			
1. Vestir-se, inclusive amarrar os cordões do sapato e abotoar suas roupas?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
2. Lavar sua cabeça e cabelos?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
3. Levantar-se de maneira ereta de uma cadeira de encosto reto e sem braço?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
4. Deitar-se e levantar-se da cama?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
5. Cortar um pedaço de carne?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
6. Levar à boca um copo de água ou xícara de café?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
7. Abrir um saco de leite comum?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
8. Caminhar em lugares planos?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
9. Subir cinco degraus?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
10. Lavar e secar o corpo após o banho?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
11. Tomar banho de chuveiro?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
12. Sentar-se e levantar-se de um vaso sanitário?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
13. Levantar os braços e pegar um objeto de aproximadamente 2,5kg que está posicionado pouco acima da cabeça?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
14. Curvar-se para pegar suas roupas no chão?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
15. Segurar-se em pé no ônibus ou no metrô?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
16. Abrir potes ou vidros de conservas que tenham sido abertos antes?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
17. Abrir e fechar torneiras?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
18. Fazer compras nas redondezas onde mora?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
19. Entrar e sair de um ônibus?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
20. Realizar tarefas tais como usar a vassoura para varrer e o rodo para água?	<input type="radio"/> nenhuma	<input type="radio"/> alguma	<input type="radio"/> muita	<input type="radio"/> incapaz
Observação				

APÊNDICE J – Campo de Câncer de mama do LAPOGEDb

HISTÓRICO DA DOENÇA	
Histórico (versão do participante)	<input type="text"/>
Histórico (do prontuário)	<input type="text"/>
Data do diagnóstico	<input type="text"/>
Origem da amostra	<input type="text"/>
Nome do médico	<input type="text"/> manter em sigilo
Número do prontuário	<input type="text"/> manter em sigilo
Classificação no prontuário	<input type="checkbox"/> Hiperplasiar <input type="checkbox"/> Cisto <input type="checkbox"/> Displasia <input type="checkbox"/> Fibroadenoma <input type="checkbox"/> Lipoma <input type="checkbox"/> Outro (descrever nas observações)

INFORMAÇÕES DE PRONTUÁRIO DO TUMOR	
Classificação (PAAF/ Core Biopsy)	<input type="checkbox"/> Carcinoma ductal <input type="checkbox"/> invasivo <input type="checkbox"/> Microcalcificações <input type="checkbox"/> doença de Paget <input type="checkbox"/> carcinoma lobular <input type="checkbox"/> in situ <input type="checkbox"/> Outro (descrever nas observações)
Classificação (biópsia pos cirurgia)	<input type="checkbox"/> Carcinoma ductal <input type="checkbox"/> invasivo <input type="checkbox"/> Microcalcificações <input type="checkbox"/> doença de Paget <input type="checkbox"/> carcinoma lobular <input type="checkbox"/> in situ <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

INFORMAÇÕES POSTERIORES À CIRURGIA	
Tamanho do tumor	<input type="text"/>
Grau de Elston	<input type="text"/>
Tamanho do tumor	<input type="text"/>
Classificação molecular	<input type="text"/>
Observação	<input type="text"/>

APÊNDICE K – Campo de Lupus do LAPOGEDb

HISTÓRICO DA DOENÇA	
Histórico (versão do participante)	<input type="text"/>
Histórico (do prontuário)	<input type="text"/>
Número do prontuário	<input type="text"/> manter em sigilo
Nome do Médico	<input type="text"/>
Classificação na primeira consulta	<input type="radio"/> Ausente <input type="radio"/> Leve <input type="radio"/> Moderada <input type="radio"/> Grave <input type="radio"/> Muito grave
Data do diagnóstico	<input type="text"/>

MANIFESTAÇÕES INICIAIS	
Manifestações iniciais	<input type="checkbox"/> Febre <input type="checkbox"/> Fenômeno de Raynaud <input type="checkbox"/> Rash discóide <input type="checkbox"/> Artrite reumatoide <input type="checkbox"/> Úlcera Oral <input type="checkbox"/> Distúrbio renal <input type="checkbox"/> Distúrbio hematológico <input type="checkbox"/> Vasculite <input type="checkbox"/> Alopecia <input type="checkbox"/> Rash malar <input type="checkbox"/> Fadiga <input type="checkbox"/> Fotossensibilidade <input type="checkbox"/> Serosite <input type="checkbox"/> Distúrbio neurológico <input type="checkbox"/> Hipertensão arterial <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

EVOLUÇÃO DA DOENÇA			
Internações	<input type="radio"/> Não <input type="radio"/> Sim	Quantas?	Motivos
Manifestações Associadas	<input type="radio"/> Não <input type="radio"/> Sim	Quais?	<input type="text"/>
Sobreposição com outra doença reumatológica?	<input type="checkbox"/> Esclerodermia <input type="checkbox"/> Sjögren <input type="checkbox"/> Dermatomiosite <input type="checkbox"/> Síndrome do Anticorpo Fosfolípido (SAF) <input type="checkbox"/> Fibrinomiálgia <input type="checkbox"/> Outro (descrever nas observações)		
Observação	<input type="text"/>		

SINTOMATOLOGIA ATUAL

Manifestações Iniciais	<input type="checkbox"/> Febre <input type="checkbox"/> Fenômeno de Raynaud <input type="checkbox"/> Rash discóide <input type="checkbox"/> Artrite reumatóide <input type="checkbox"/> Úlcera Oral <input type="checkbox"/> Distúrbio renal <input type="checkbox"/> Distúrbio hematológico <input type="checkbox"/> Vasculite <input type="checkbox"/> Alopecia <input type="checkbox"/> Rash malar <input type="checkbox"/> Fadiga <input type="checkbox"/> Fotosensibilidade <input type="checkbox"/> Seroite <input type="checkbox"/> Distúrbio neurológico <input type="checkbox"/> Hipertensão arterial <input type="checkbox"/> Outro (descrever nas observações)
	Observação

TRATAMENTO

Corticosteroide	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Azatioprina	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Mefloretato	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Clidopirina	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Clidofamida	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
AINE	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Anfimetaco	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Anafetaco	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência
Anafetaco	<input type="radio"/> Nunca usou			
	<input type="radio"/> Usou, mas parou	Motivo da interrupção?	<input type="text"/>	
	<input type="radio"/> Usa	Qual?	Dose	Frequência

APÊNDICE L – Campo de psoríase do LAPOGEDb

HISTÓRICO DA DOENÇA	
Histórico (versão do participante)	<input type="text"/>
Histórico (do prontuário)	<input type="text"/>
Número do prontuário	<input type="text"/> manter em sigilo
Fator desencadeante	<input type="text"/>
Médico	<input type="text"/> manter em sigilo
Local das primeiras lesões	<input type="text"/>
Classificação na primeira consulta	<input type="radio"/> Ausente <input type="radio"/> Leve <input type="radio"/> Moderada <input type="radio"/> Grave <input type="radio"/> Muito grave
Classificação PASI	<input type="text"/>
Data do diagnóstico	<input type="text"/>

PADRÃO MORFOLÓGICO	
Classificação	<input type="checkbox"/> vulgar <input type="checkbox"/> gutata <input type="checkbox"/> pustulosa <input type="checkbox"/> eritodérmica <input type="checkbox"/> palmoplantar <input type="checkbox"/> invertida <input type="checkbox"/> Outro (descrever nas observações)
Observação	<input type="text"/>

PAINEL LABORATORIAL		
Hemograma com plaquetas	<input type="text"/>	unidade?
ALTA/AST	<input type="text"/>	unidade?
gamaGST	<input type="text"/>	unidade?
Fosfatase alcalina	<input type="text"/>	unidade?
Uréia	<input type="text"/>	unidade?
Creatinina	<input type="text"/>	unidade?
Ácido úrico	<input type="text"/>	unidade?
Colesterol total	<input type="text"/>	unidade?
HDL	<input type="text"/>	unidade?
Triglicérides	<input type="text"/>	unidade?
Glicose	<input type="text"/>	unidade?
Na, K, Ca, Mg	<input type="text"/>	unidade?

VHS	<input type="text"/>	unidade?
Proteína Creatina	<input type="text"/>	unidade?
ASLO	<input type="text"/>	unidade?
HBsAg	<input type="text"/>	unidade?
Anti-HBs	<input type="text"/>	unidade?
Anti-HCV	<input type="text"/>	unidade?
Anti HIV 1 e 2	<input type="text"/>	unidade?
FAN	<input type="text"/>	unidade?
IgE total	<input type="text"/>	unidade?
Observação	<input type="text"/>	

TRATAMENTO

Corticosteroídes	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Metotrexato	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Sulfassalozena	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Antimalárico	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Ciclofosfamida	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Infliximab	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Etanercept	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Aine	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Analgésicos	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Outros	<input type="radio"/> Não	<input type="radio"/> Sim	Qual? <input type="text"/>	Dose <input type="text"/>	Frequência	<input type="text" value="Não se aplica"/>
Observação	<input type="text"/>					