



**UNIVERSIDADE FEDERAL DO VALE DO SÃO FRANCISCO
CURSO DE GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS**

ELIANDERSON GOMES DE SÁ

**ESTUDO EVOLUTIVO DO CROMOSSOMO SEXUAL Y DE
Callithrix jacchus (LINNAEUS, 1758): UMA ANÁLISE
COMPARATIVA FRENTE A OUTROS PRIMATAS POR
MÉTODOS *in silico***

Petrolina

2018

ELIANDERSON GOMES DE SÁ

**ESTUDO EVOLUTIVO DO CROMOSSOMO SEXUAL Y DE
Callithrix jacchus (LINNAEUS, 1758): UMA ANÁLISE
COMPARATIVA FRENTE A OUTROS PRIMATAS POR
MÉTODOS *in silico***

Trabalho de Conclusão de Curso apresentado ao Colegiado de Ciências Biológicas da Universidade Federal do Vale do São Francisco, *Campus* Ciências Agrárias, como requisito parcial para a obtenção do grau de Bacharel em Ciências Biológicas.

Orientador: Prof^a Dr^a Kyria Cilene de Andrade Bortoleti

Co-orientador: Dr^o João Pacifico Bezerra Neto

Petrolina

2018

S111e Sá, Elianderson Gomes de
Estudo evolutivo do cromossomo sexual Y de *Callithrix jacchus* (Linnaeus, 1758): uma análise comparativa frente a outros primatas por métodos *in silico* / Elianderson gomes de sá. -- Petrolina, 2018.
XIV, 64f.: il. ; 29 cm.

Trabalho de Conclusão de Curso (Graduação em Ciências Biológicas) - Universidade Federal do Vale do São Francisco, Campus Ciências Agrárias, Petrolina, 2018.

Orientadora: Prof^a. Dr^a. Kyria Cilene de Andrade Bortoleti.

Referências.

1. Saguí. 2. Sintenia. 3. Bioinformática. I. Título. II. Universidade Federal do Vale do São Francisco

CDD 575

Ficha catalográfica elaborada pelo Sistema Integrado de Biblioteca SIBI/UNIVASF
Bibliotecária: Ana Cleide Lucio CRB – 4 /2064

UNIVERSIDADE FEDERAL DO VALE DO SÃO FRANCISCO
CURSO DE GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

FOLHA DE APROVAÇÃO

Elianderson Gomes de Sá

ESTUDO EVOLUTIVO DO CROMOSSOMO SEXUAL Y DE *Callithrix jacchus*
(LINNAEUS, 1758): UMA ANÁLISE COMPARATIVA FRENTE A OUTROS
PRIMATAS POR MÉTODOS *IN SILICO*

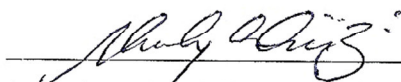
Trabalho de Conclusão de Curso apresentado ao
Colegiado de Ciências Biológicas da Universidade
Federal do Vale do São Francisco, *Campus*
Ciências Agrárias, como requisito parcial para a
obtenção do grau de Bacharel em Ciências
Biológicas.

Aprovado em: 29 de maio de 2018.

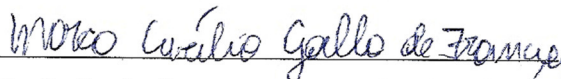
Banca Examinadora



Dr^a. Kyria Cilene de Andrade Bortoleti - Orientadora, UNIVASF



Dr^a. Michely Correia Diniz – Primeira Examinadora, UNIVASF



Dr^o Marco Aurélio Gallo de França – Segundo Examinador, UNIVASF

AGRADECIMENTOS

Primeiramente, agradeço a Deus e aos meus familiares por terem feito tudo o possível para permitir minha chegada até aqui. E em seguida, agradeço aos meus colegas de trabalho e faculdade:

A professora, por ter me dado liberdade de fazer algo que gosto e por ter tido dedicação, carinho e nos tratar sempre com respeito. Também por ter nos dado uma ínfima parte de seu conhecimento, o qual fez nos apaixonarmos pela genética, e por ser exemplo de profissional a seguir. Obrigado professora Kyria! A senhora é um exemplo que tive a sorte de ter.

Ao professor João Pacifico, por ter projetado todo o trabalho a partir de uma ideia simples, por isso, obrigado por ter me encorajado e, apesar de seu grande conhecimento, atuar sempre com humildade e respeito.

A Palloma, por ter sido a melhor companheira do lab, por ter dividido os momentos bons comigo (e os ruins também). Por ter se preocupado mesmo sem necessidade (quem não sabe pegar ônibus em Recife?!), me tirado dúvidas e ter cuidado de mim. Palloma, espero ter sido um bom amigo e ter ajudado também, guardo você tem um lugar especial no meu coração. Você tem esse amigo para sempre! E nunca se esqueça —uma mão lava a outra e as duas lavam o rosto!.

A Araúna e prima Cinthia, também por cuidarem de mim, dividir as comidas na viagem (não cozinho mais! A próxima são vocês), por contribuírem com o meu conhecimento e brigar comigo também (mais Araúna, porque ela é chata! Brincadeira, te amo Araúna!). Prima Cinthia, obrigado por ter sido melhor prima e me acompanhar na aula do “Homi” (você sabe quem).

Laysla, obrigado por quebrar meu dedo e ter me iludido com um hambúrguer, brincadeira! Obrigado por organizar o lab e pelo cuidado que você tem com todos. Ingrid, obrigado pelo carinho e sua linda voz, ainda quero ser o seu produtor. Deborah e Ilka, obrigado pelos conselhos e pelo exemplo também. Deborah, obrigado pelas conversas, amizade e paciência (não que você precisasse). Ilka, você sempre gostei de te escutar e, não posso esquecer, agradeço pela camisa UV.

Obrigado Bianca, Naiara, Inês e Layze pelos melhores momentos da faculdade e poder sempre contar com vocês. Agradeço também aos amigos de classe, Monique, Lilian, Khatianne e Danillo. Sempre os terei no coração, você são meus irmãos de classe.

RESUMO

O *Callithrix jacchus* (Linnaeus, 1758), também conhecido como sagui-comum, é um pequeno primata da família Callitrichidae, endêmico do nordeste brasileiro e distribuído em regiões do Cerrado, Floresta Atlântica e Caatinga. Em 2014, o *C. jacchus*, espécie bastante utilizada em pesquisas biomédicas, teve seu genoma totalmente sequenciado (*The Marmoset Genome Sequencing and Analysis Consortium*), possibilitando estudos *in silico* de genômica comparativa entre a espécie com macacos do Novo e Velho Mundo. Neste sentido, considerando que os cromossomos sexuais apresentam uma evolução independente, o presente trabalho realizou uma análise comparativa *in silico* do mapa genético do cromossomo sexual Y de *C. jacchus* frente a outros primatas, cujos genomas foram sequenciados e disponibilizados, identificando relações macrossintênicas dentro deste grupo. Para isso, a RefSeq NC_013919.1 referente ao cromossomo Y de *C. jacchus* (versão genômica *Callithrix jacchus*-3.2.1; <https://www.ncbi.nlm.nih.gov/genome>) foi avaliada, anotando e selecionando-se os genes vinculados ao *Ensembl Release 90* (<http://www.ensembl.org>). Em seguida, uma busca por órtologos foi realizada em genomas de primatas disponíveis e os mapas genéticos foram construídos para cromossomos de *C. jacchus*, *Homo sapiens*, *Pongo abelii*, *Macaca mulata*, *Chlorocebus sabaeus* e *Sus scrofa*, mediante a ferramenta online MapGene2Chrom web v2 (http://mg2c.iask.in/mg2c_v2.0/), com o intuito de observar possíveis relações macrossintênicas entre autossomos e cromossomos sexuais das citadas espécies. O número de 41 genes foram anotados para o cromossomo Y de *C. jacchus*, o qual se mostrou o menor em extensão de bases e com menor quantidade de genes entre as espécies analisadas. Dentre esses genes, vinte e dois possuíam dados referentes a sequências ortólogas, cujas localizações e ordens nos cromossomos sexuais e em diferentes autossomos das espécies estudadas enfatizaram a presença de regiões macrossintênicas, quebras de colinearidade e rearranjos cromossômicos entre estes dois grupos cromossômicos. Evidências de inversões e translocações entre um cromossomo autossomo e o cromossomo Y de *C. jacchus*, ao longo do tempo evolutivo, são discutidas.

Palavras-chave: sagui, sintenia, rearranjos cromossômicos, bioinformática, genômica comparativa.

ABSTRACT

The *Callithrix jacchus* (Linnaeus, 1758), also known as the common-marmoset, is a small primate of the Callitrichidae family, endemic of the Brazilian Northeast and distributed in the Cerrado, Atlantic Forest and Caatinga regions. In 2014, the *C. jacchus*, a species widely used in biomedical research, had its genome entirely sequenced (*The Marmoset Genome Sequencing and Analysis Consortium*), enabling studies *in silico* of comparative genomics between the species with New and Old World monkeys. In this sense, considering that the sexual chromosomes present an independent evolution, the present study performed an *in silico* comparative analysis of the genetic map of the sexual chromosome Y of *C. jacchus* compared to other primates, whose genomes were sequenced and made available, identifying syntenic relationships within this group. For this, the RefSeq NC_013919.1 referring to the Y chromosome of *C. jacchus* (*Callithrix jacchus*-3.2.1 genomic version; <https://www.ncbi.nlm.nih.gov/genome>) was evaluated, annotating and selecting the genes linked to *Ensembl Release 90* (<http://www.ensembl.org>). Then, a research for orthologs was performed on genomes of available primates and the genetic maps were constructed for chromosomes of *C. jacchus*, *Homo sapiens*, *Pongo abelii*, *Macaca mulata*, *Chlorocebus sabaeus* and *Sus scrofa*, through the MapGene2Chrom web v2 *online* tool (http://mg2c.iask.in/mg2c_v2.0/), in order to observe possible macrosyntenic relationships between autosomes and sexual chromosomes of the mentioned species. The number of 41 genes were annotated for the Y chromosome of *C. jacchus*, which was the smallest in extension of bases and with fewer genes among the analyzed species. Among these genes, twenty-two had orthologous sequence data, whose locations and orders in the sexual chromosomes and in different autosomes of the species studied emphasized the presence of macrosyntenic regions, collinearity breaks and chromosomal rearrangements between these two chromosomal groups. Evidence of inversions and translocations between an autosomal chromosome and the Y chromosome of *C. jacchus*, throughout evolutionary time, are discussed.

Key words: marmoset, synteny, chromosomal rearrangements, bioinformatics, comparative genomics.

LISTA DE ABREVIATURAS E SIGLAS

2n	Número diploide
A	Adenina
BACs	<i>Bacterial Artificial Chromosomes</i> Cromossomos Artificiais de Bactérias
C	Citosina
Cjac	<i>Callithrix jacchus</i>
Csab	<i>Chlorocebus sabaues</i>
FISH	<i>Fluorescence in situ Hybridization</i> Hibridização <i>in situ</i> Fluorescente
G	Guanina
Gb	Gigabase
Hsap	<i>Homo sapiens</i>
Mmul	<i>Macaca mulata</i>
MSY	Região Macho-específica
NCBI	<i>National Center for Biotechnology Information</i> Centro Nacional de Informações Biotecnológicas
NGS	<i>Next Generation Sequencing</i> Sequenciamento de Próxima Geração
Pabe	<i>Pongo abelii</i>
PARs	Regiões Pseudoautosômicas
Ptro	<i>Pan troglodytes</i>
Sanger	Tecnologia de primeira geração
Sscr	<i>Sus scrofa</i>
T	Timina
WGS	<i>Whole Genome Sequencing</i> Sequenciamento de Genoma Inteiro

LISTA DE TABELAS

Tabela 1 – Espécies analisadas, identificação das sequências de referência (RefSeq) para os cromossomos Y destas espécies, bem como a sua extensão em bases e quantidade de genes anotados, segundo o NCBI (acesso em 30 de outubro de 2017)..	30
---	----

LISTA DE ILUSTRAÇÕES

- Figura 1** – A sistemática dos primatas, conforme descrito por Martin (2012). 17
- Figura 2** - Mapa genético do cromossomo Y de *C. jacchus* contendo 41 genes e construído a partir de informações disponíveis no NCBI, acesso em 30 de outubro de 2017, mediante a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....31
- Figura 3** - Comparação do tamanho em bases de extensão dos cromossomos Y das espécies em estudo, *Callithrix jacchus* (Cjac), *Homo sapiens* (Hsap), *Macaca mulatta* (Mmul), *Chlorocebus sabaues* (Csab) e *Sus scrofa* (Sscr), os quais foram construídos mediante a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....32
- Figura 4** - Mapa genético do cromossomo Y de *C. jacchus* contendo 28 genes segundo o *Ensembl release* 90. Tal mapa foi construído utilizando-se a ferramenta *online* MapGene2Chrom. Em destaque, estão os genes sem ortologia. Fonte: próprio autor.....34
- Figura 5** - Mapa genético dos cromossomos de *Homo sapiens* (Hsap) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release* 90. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....35
- Figura 6** - Mapa genético dos cromossomos de *Pongo abelii* (Pabe) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release* 90. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....36
- Figura 7** - Mapa genético dos cromossomos de *Macaca mulatta* (Mmul) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release* 90. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....37
- Figura 8** - Mapa genético dos cromossomos de *Chlorocebus sabaues* (Csab) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release* 90. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....38
- Figura 9** - Mapa genético dos cromossomos de *Sus scrofa* (Sscr) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release* 90. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.....38

- Figura 10** - Ordem dos genes ortólogos comuns as espécies *Callithrix jacchus* (Cjac), *Homo sapiens* (Hsap), *Pongo abelli* (Pabe), *Macaca mulatta* (Mmul) e *Chlorocebeus sabaesus* (Csab), conforme as suas respectivas localizações disponíveis no *Ensembl Release 90*, mostrando quebra de colinearidade. Fonte: próprio autor.....39
- Figura 11** – Possíveis rearranjos cromossômicos ocorridos envolvendo os genes ortólogos comuns presentes nos cromossomos Y de *Callithrix jacchus* (Cjac), 15 de *Homo sapiens* (Hsap) e *Pongo abelli* (Pabe), 7 de *Macaca mulatta* (Mmul) e 26 de *Chlorocebeus sabaesus* (Csab), levando-se em consideração a filogenia proposta por PERELMAN (2011). Fonte: próprio autor.....40
- Figura 12** – Primeira representação hipotética de rearranjos cromossômicos entre um cromossomo autossomo ancestral e o cromossomo Y de *Callithrix jacchus*, com destaque para uma translocação seguida de inversão, resultando na ordem dos genes observadas neste cromossomo sexual. Fonte: próprio autor.....41
- Figura 13** – Segunda representação hipotética de rearranjos cromossômicos entre um cromossomo autossomo ancestral e o cromossomo Y de *Callithrix jacchus*, com destaque para uma inversão no autossomo e, posterior, translocação para o cromossomo sexual, resultando na ordem dos genes observadas no cromossomo Y. Fonte: próprio autor.....41
- Figura 14** – Representação hipotética de uma inversão no cromossomo autossomo 26 de *Chlorocebeus sabaesus*, enfatizando a quebra de colinearidade frente a ordem dos genes em questão visualizadas em cromossomos autossomos de *Homo sapiens*, *Pongo abelii* e *Macaca mulata*. Fonte: próprio autor.....42

SUMÁRIO

1. INTRODUÇÃO	13
2. OBJETIVOS	155
2.1 OBJETIVO GERAL.....	155
2.2 OBJETIVOS ESPECÍFICOS.....	155
3. REFERENCIAL TEÓRICO	166
3.1 SISTEMÁTICA E TAXONOMIA DOS PRIMATAS	166
3.2 O GÊNERO <i>Callithrix</i> E A ESPÉCIE <i>Callithrix jacchus</i>	19
3.3 TECNOLOGIA DE SEQUENCIAMENTO PARA ESTUDOS GENÔMICOS	20
3.4 SEQUENCIAMENTO DOS GENOMAS E SUAS APLICAÇÕES EM ESTUDOS EVOLUTIVOS.....	23
3.5 SEQUENCIAMENTO GENÔMICO DE <i>Callithrix jacchus</i>	255
3.6 IMPORTÂNCIA DA BIOINFORMÁTICA PARA ANÁLISE DE DADOS BIOLÓGICOS	266
4. MATERIAL E MÉTODOS	288
4.1 MINERAÇÃO DE DADOS PARA AS ANÁLISES DE MACROSSINTENIA ...	288
4.2 CONSTRUÇÃO DOS MAPAS GENÉTICOS E ANÁLISES DE MACROSSINTÊNICAS.....	299
5. RESULTADOS E DISCUSSÃO.....	300
5.1 CARACTERIZAÇÃO DO CROMOSSOMO Y	30
5.2 ANÁLISE MACROSSINTÊNICA DO CROMOSSOMO Y DE <i>Callithrix jacchus</i> FRENTE OUTROS PRIMATAS.....	344
6. CONSIDERAÇÕES FINAIS	444
REFERENCIAS BIBLIOGRÁFICAS ¹	455
APÊNDICE A- QUADROS DOS GENES	511

1. INTRODUÇÃO

O *Callithrix jacchus* (Linnaeus, 1758) (Callitrichidae), também conhecido como sagui-do-nordeste ou sagui-comum, é um pequeno primata endêmico do nordeste brasileiro, distribuído em regiões do Cerrado, Mata Atlântica e Caatinga (RYLANDS; COIMBRA-FILHO; MITTERMEIER, 1993). Em termos de pesquisas laboratoriais, devido as vantagens que esta espécie apresenta em relação a outros primatas, como a taxa de fertilidade acima da média e o pequeno tamanho corporal que reflete em baixos custos de alimentação e economia de espaço, o *C. jacchus* tem sido bastante utilizado em pesquisas biomédicas na América do Norte e Europa, principalmente, devido a sua suscetibilidade a um grande número de agentes infecciosos do ser humano (MANSFIELD, 2003).

Em 2014, *Callithrix jacchus* teve seu genoma totalmente sequenciado; cerca de 2.26 Gb de informação genética de um indivíduo fêmea com quimerismo foi obtido, permitindo a identificação de características específicas que podem contribuir para a biologia dos primatas de pequeno porte, como modificações genéticas que influenciam no tamanho do corpo. Associadamente, também tem possibilitado comparações com outros genomas já disponíveis, por exemplo, os dos Catarrhines (macacos do Velho Mundo, humanos e outros homínídeos), evidenciando novos aspectos da evolução genômica dessas espécies (*The Marmoset Genome Sequencing and Analysis Consortium*, 2014).

Neste contexto, o sequenciamento completo do genoma de uma espécie disponibiliza uma grande quantidade de dados para estudos de organização gênica e comparativos, esclarecendo os rearranjos que direcionaram o genoma de uma dada espécie (BOURQUE, 2005). Logo, os estudos genômicos comparativos tornaram-se uma excelente ferramenta, não só para entender as relações evolutivas entre as espécies, mas também para descobrir novos genes e suas funções (KIM, 2007). Apesar disso, poucas espécies possuem seu genoma disponibilizado por sequenciamento; por exemplo, apenas duas espécies de Platyrrhini (macacos do Novo Mundo) tiveram seus projetos genoma completado, o *Callithrix jacchus* e o *Saimiri boliviensis* (I. Geoffroy e Blainville, 1834) (macaco-esquilo), mas somente o *C. jacchus* teve o genoma apresentado em forma de cromossomos (CAPOZZI, 2016), motivo pelo qual *S. boliviensis* não foi utilizado nesta análise.

Comparativamente a outros vertebrados, os primatas apresentam uma grande variabilidade cariotípica, podendo ser encontrados conjuntos diploides variando de $2n=6$ até $2n=102$, o que acaba mascarando as estruturas conservadas do genoma entre as espécies, mas que já tem sido visualizadas através do mapeamento genético e das técnicas de pintura cromossômica (CONTRERAS, 1990; KEMKEMER et al., 2009; WURSTER, 1970). O mapeamento genético de um organismo fornece um guia para os experimentos de sequenciamento, indicando a exata posição dos genes, enquanto as técnicas de caracterização cromossômica permitem a visualização dos segmentos sintênicos, rearranjos cromossômicos e de outras características distintas dos cromossomos nos genomas das espécies sob investigação (DIXIT, 2014; KEMKEMER, 2009). Os segmentos sintênicos podem ser utilizados como sequências-chaves para a construção de genomas ancestrais, bem como permitem inferir sobre os mecanismos de evolução cariotípica (KEMKEMER, 2009).

Estudos de sintenia e de rearranjos cromossômicos começam com a identificação de regiões ortólogas em diferentes genomas; uma vez encontradas, essas regiões são denominadas blocos sintênicos (*syntenic blocks*), nos quais são exploradas as possíveis alterações espaciais, em busca do cenário mais parcimonioso que explique os rearranjos que levaram a atual disposição do genoma (BOURQUE, 2005).

Segundo Murphy (2001), o desenvolvimento e aperfeiçoamento das análises genômico-comparativas, associando os conhecimentos de análises cromossômicas com os de mapeamento genético em um maior número de espécies, têm permitido a estimativa da taxa de mutações cromossômicas, bem como uma reconstrução mais precisa de um genoma ancestral. Historicamente, os cromossomos sexuais apresentam trocas limitadas de material genético com o restante do complemento cromossômico, o que permite uma análise isolada e mais detalhada de seus rearranjos durante a evolução (BOURQUE, 2005). Em vista disso, esse trabalho teve como objetivo analisar os rearranjos cromossômicos ocorridos no cromossomo sexual Y de *C. jacchus*, mediante um estudo comparativo do mapa genético entre diferentes espécies de primatas com genomas totalmente sequenciado, utilizando ferramentas de bioinformática.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Realizar uma análise comparativa *in silico* do mapa genético do cromossomo sexual Y de *Callithrix jacchus* frente a outros primatas, inferindo as relações macrossintênias dentro deste grupo.

2.2 OBJETIVOS ESPECÍFICOS

- Anotar e caracterizar os genes localizados no cromossomo sexual Y de *C. jacchus*, a partir das informações disponíveis nos bancos de dados NCBI (*National Center for Biotechnology Information*; Centro Nacional para Informação Biotecnológica) e *Ensembl Release 90*;
- Catalogar os genes ortólogos encontrados no cromossomo sexual Y de *C. jacchus* relacionados a outros primatas, cujas informações estejam disponíveis no *Genome Data Viewer* do NCBI;
- Construir e comparar mapas genéticos para os cromossomos sexuais das espécies em estudo, com o intuito de identificar regiões sintênicas entre as mesmas;
- Identificar rearranjos cromossômicos envolvidos na evolução cromossômica da espécie *C. jacchus*.

3. REFERENCIAL TEÓRICO

3.1 SISTEMÁTICA E TAXONOMIA DOS PRIMATAS

Devido ao desenvolvimento da pesquisa biomédica moderna, os primatas têm sido cada vez mais utilizados como animais modelos para estudos laboratoriais na busca de curas para doenças humanas (TORRES, 2010). Além disso, devido a proximidade filogenética, o grande interesse do ser humano em descobrir suas próprias origens tem refletido diretamente no maior número de pesquisas sobre biologia de primatas, incluindo morfologia, fisiologia, comportamento, ecologia, genética e genômica (MARTIN, 2012).

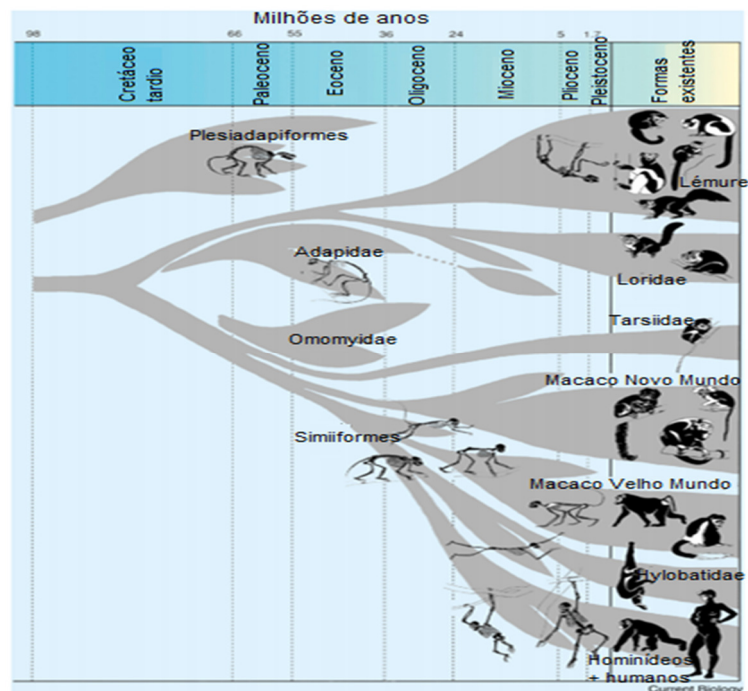
Neste raciocínio, uma das questões mais discutidas é a origem dos primatas. Os 'Euprimates' ou "Primatas do aspecto moderno", parentes fósseis diretos reconhecíveis dos primatas como exemplo o fóssil de Adapidae e Simiiformes, datam do Eoceno mais antigo, cerca de 55 milhões de anos atrás (MARTIN, 2012) (Figura 1). Além dos "Euprimates", dentro da ordem dos Primatas, um outro grupo de mamíferos mais antigos era considerado, os Plesiadapiformes ou "Primatas arcaicos" (Figura 1). No entanto, o consenso é que os primatas arcaicos divergiram antes da radiação adaptativa que originou os primatas existentes e seus parentes fósseis diretos (MARTIN, 2012), não havendo evidências claras que os Plesiadapiformes são o grupo irmão mais próximo dos "Euprimates" (GODINOT, 2017).

Entretanto, dados morfológicos e moleculares apontam que os primatas é um grupo monofilético com cerca de 67-69 gêneros, o qual divergiu no Cretáceo superior há cerca de 98 milhões de anos. Com a expansão do Eoceno, formaram-se as cinco principais linhagens existentes, as quais embasam todas as classificações e árvores evolutivas deste grupo: Lemuriformes (Gregory, 1915), Lorisiformes (Gregory, 1915), Tarssiformes (Gregory, 1915), Platyrrhini (E. Geoffroy, 1812) e Catarrhini (E. Geoffroy, 1812), todas apontadas como monofiléticas (Figura 1) (MARTIN, 2012; PERELMAN, 2011).

A infraordem Lemuriformes é constituída pelo grupo dos lêmures, únicos primatas presentes na Ilha de Madagascar, onde sofreram a principal radiação adaptativa levando a cerca de 80 espécies. Por sua vez, a infraordem Lorisiformes é representada pelos lóris (Gray, 1821), *pottos* (Statius Müller, 1766) e gálagos (Gray,

1825), os quais apresentam uma distribuição geográfica mais ampla que os Lemuriformes, ocorrendo na África e Ásia (Sul e Sudeste). Com cerca de 30 espécies, para este grupo, a radiação adaptativa foi mais limitada. O terceiro grupo, a infraordem Tarsiiformes, é composta por doze espécies de *tarsius* (Storr, 1780) distribuídas nos arquipélagos do sudeste asiático (MARTIN, 2012).

Figura 1 – A sistemática dos primatas, conforme descrito por Martin (2012).



Os Platyrrhini constituem os chamados macacos do Novo Mundo, os quais são encontrados na América Central e do Sul onde se diversificaram amplamente, havendo cerca de 130 espécies reconhecidas. Este representa um grupo de macacos pequenos e com garras, a exemplo dos saguis e *tamarins* (Gray, 1870). Por sua vez, os Catarrhini são representados pelos macacos do Velho Mundo, grandes macacos e humanos (Gray, 1825), sendo o maior grupo de primatas com 150 espécies distribuídas na África, Sul e Sudeste asiático (MARTIN, 2012).

Por apresentarem uma maior proporção de características primitivas, os Lemuriformes, Lorisiformes e Tarsiiformes são frequente e coletivamente chamados de *Promisians* (Illiger, 1811) (Promisianos), “Pré-macacos” ou “Primatas inferiores”. São animais relativamente pequenos com peso corporal em torno de 500 g e, a maioria, com hábitos noturnos. Em contraste, os Anthropoidea, “Primatas superiores”

ou *Simians* (Siminianos) (Haeckel, 1866), representados pelos macacos do Velho e Novo Mundo (Haeckel, 1866), grandes macacos e humanos, compartilham um conjunto distinto de características anatômicas em seus dentes, mandíbulas, cérebro e sistema reprodutivo (DIXSON, 2015; MARTIN, 2012). Com exceção do macaco coruja do Novo Mundo, todos os Siminianos são noturnos e apresentam corpos maiores com um peso modal de 5 Kg (MARTIN, 2012). Vale ressaltar que, entre os Siminianos, os macacos do Novo Mundo e do Velho Mundo possuem muitas similaridades morfológicas e comportamentais em comum, principalmente devido aos seus modos de vida semelhantes e a evolução convergente, vivendo em grupos com organizações sociais complexas e diversos sistemas de acasalamento (DIXSON, 2015).

Uma questão a ser destacada refere-se aos Tarsiiformes; estes animais mostram similaridades aos “Primatas superiores” em relação a visão, olfato, morfologia cerebral e reprodução, fato que sugere uma ligação ancestral entre esses grupos taxonômicos a qual tem sido confirmada por evidências morfológicas e moleculares. Os lêmures e lóris mostram uma área de pele nebulosa e úmida em torno das narinas (*rinarium*), enquanto que os Tarsiiformes e Siminianos são os únicos mamíferos que perderam esta característica, sendo classificados nas subordens Strepsirrhini (E. Geoffroy, 1812) e Haplorrhini (Pocock, 1918), respectivamente. Ou seja, os Lemuriformes e Lorisiformes são grupos irmãos, enquanto que Tarsiiformes é um grupo irmão de Simiiformes (DIXSON, 2015; MARTIN, 2012; PERELMAN, 2011).

De acordo com Perelman (2011), considerando a filogenia construída a partir de dados moleculares de 186 primatas representantes de 61 gêneros, as subordens Strepsirrhini e Haplorrhini divergiram de seu ancestral há 87 milhões de anos e a ordem de mamíferos mais próxima aos primatas seria Dermoptera, seguida de Scadentia. Dentro de Strepsirrhini, os dados afirmam que há 68,7 milhões de anos ocorreu a divergência dos progenitores de Lemuriformes/Chiromyiformes e Lorisiformes, cujas origens são datadas em 58,6 e 40,3 milhões de anos, respectivamente (PERELMAN, 2011). Em relação aos Haplorrhini, estudos afirmam que os Platyrrhini divergiram dos Catarrhini há 43,5 milhões de anos durante o Eoceno. Os macacos do Novo Mundo foram englobados em três famílias Pitheciidae (Mivart, 1865) (uacari, *titis* e *sakis*), Atelidae (búgio, macaco-aranha, macaco-barrigudo e muriquis) e Cebidae (Bonaparte, 1831) (saguim, *tamarins* e macacos noturnos), as quais divergiram há 24,8 milhões de ano (PERELMAN, 2011).

Por sua vez, na família Cebidae, a radiação teve início há 20 milhões de anos com a emergência dos clados irmãos *Cebus* (Erxleben, 1777) (Cebinae) e *Saimiri* (Voigt, 1831) (Saimirinae). Após 700.000 anos, outra divergência foi notada levando à diferenciação de *Aotus* (Hershkovitz, 1983) e Callitrichinae (Gray, 1821) (saguins, *tamarins*). Em Cebidae (Bonaparte, 1831), um caso de adaptação denominado (*phyletic dwarfism*), gradiente no tamanho morfológico parcialmente correlacionado com o tempo evolutivo foi relatado. De acordo com este termo, as espécies de *Aotus*, *Cebus* e *Saimiri* são maiores que os representantes de Callitrichinae, entre eles o gênero *Callithrix*, o qual juntamente com *Callimico*, *Mico* e *Leontopithecus*, constituem as menores espécies de macacos no Novo Mundo (PERELMAN, 2011).

3.2 O GÊNERO *Callithrix* E A ESPÉCIE *Callithrix jacchus*

Os Platyrrhines, comumente referidos como macacos do Novo Mundo, são primatas nativos da América do Sul e representam um grupo diverso de animais, classificados em 16 gêneros e mais de 100 espécies, cuja divergência foi estimada em cerca de 35 – 47 milhões de anos atrás a partir dos Catarrhines (KONKEL et al., 2016; MULLER, 2006).

Dentre os seus representantes, o gênero *Callithrix* (Subfamília Callitrichinae: Família Callitrichidae) é um grupo de primatas tropicais com idade estimada em 2,5 milhões de anos, composto por seis espécies que se distribuem no território brasileiro em três biomas, Caatinga, Cerrado e Floresta Atlântica (PERELMAN et al., 2011; RYLANDS; MITTERMEIER, 2009).

Baseado em características genéticas, acústicas, morfológicas e desenvolvimentais, este gênero pode ser subdividido em dois subgrupos: grupo aurita, o qual é constituído por *C. flaviceps* e *C. aurita*; e, grupo jacchus, o qual é composto por quatro espécies remanescentes do gênero, *C. jacchus*, *C. geoffroyi*, *C. penicillata* e *C. kuhlii* (PERELMAN et al., 2011), cujas relações filogenéticas foram esclarecidas por dados de genoma mitocondrial, revelando que *C. jacchus* divergiu de *C. penicillata* há 0,70 milhões de anos (MALUKIEWICZ et al., 2017).

O *Callithrix jacchus*, popularmente conhecido como sagui-comum, é um primata arboreal que, originalmente, distribuiu-se na costa do Nordeste Brasileiro. São animais com um pequeno tamanho corporal, como resultado de seleção natural por

miniaturização, usam o sistema de cruzamento cooperativo, ou seja, em um grupo social apenas um par de adultos são reprodutores ativos, e apresentam uma tendência a gestações múltiplas, característica apontada como uma adaptação reprodutiva (HARRIS et al., 2014).

Do ponto de vista citogenético, o *C. jacchus* apresenta número cromossômico $2n=46$, constituído por 22 pares de autossomos e um par sexual (Sistema sexual XY); o cromossomo X é um submetacêntrico grande e o cromossomo Y é um acrocêntrico pequeno (BENIRSCHKE; ANDERSON; BROWNHILL, 1962).

Este sagui apresenta similaridade genética, anatômica e fisiológica em comparação aos humanos, sendo assim bastante utilizado em pesquisas biomédicas envolvendo os estudos de envelhecimento, de biologia reprodutiva, de comportamento, de doenças autoimunes, neuroendocrinologia, farmacologia, imunologia, entre outros (MITCHELL; LEOPOLD, 2015; ORSI et al., 2011).

3.3 TECNOLOGIA DE SEQUENCIAMENTO PARA ESTUDOS GENÔMICOS

O sequenciamento de nucleotídeos é uma das mais importantes ferramentas biológicas de pesquisa, com amplo espectro de aplicação. O método de sequenciamento de Sanger destacou-se nas pesquisas científicas durante 30 anos, sendo utilizado no Projeto Genoma Humano, demonstrando que o WGS (*Whole Genome Sequencing*; Sequenciamento do Genoma Inteiro) poderia ser alcançado, mas com custos muito altos para ser utilizado de forma rotineira e/ou em pequenos centros de pesquisa (GUT, 2013).

Tais limitações fomentaram a necessidade de novas tecnologias para sequenciar grandes quantidades de dados genômicos; com isso, os esforços foram direcionados para o desenvolvimento de novos métodos chamados NGS (*Next Generation Sequencing*; Sequenciamentos de Próxima Geração) (MCGINN, 2013), deixando o sequenciamento do Sanger (Tecnologia de Primeira Geração) com menos avanços relatados (METZKER, 2010).

A tecnologia NGS fornece uma alternativa muito mais barata e de alto rendimento para o sequenciamento de DNA do que o método tradicional de Sanger, uma vez que pequenos genomas podem ser sequenciados por inteiro em um único

dia, um alto rendimento que facilita a descoberta de genes e elementos regulatórios associados à doenças (GRADA, 2013). Atualmente, existem pelos menos quatro gerações de métodos para sequenciamento de DNA (GUT, 2013).

Na primeira geração, fragmentos de sequência são geradas pela extensão enzimática de um iniciador hibridado a um conjunto de moléculas modelo e introdução de terminações T, C, G ou A específicas ao longo desse modelo (SANGER; NICKLEN; COULSON, 1977). Em seguida, esses fragmentos são separadas por eletroforese em gel, no qual os produtos são revelados por marcação radioativa e a exposição a um filme ou pela introdução de corantes fluorescentes na reação de terminação e na imagem fluorescente (MCGINN, 2013).

A segunda geração é a metodologia mais difundida, principalmente por esse sequenciamento gerar muitas sequências curtas que exigem grande capacidade de processamento de análise de dados pelos computadores; mas a medida em que a migração para o sequenciamento de terceira geração ocorrer, com sequências mais longas, as demandas computacionais provavelmente diminuirão (GUT, 2013). Uma das principais características do sequenciamento de 2ª geração é o uso de vários modelos clonais em paralelo e um processo de determinação de sequência usando replicação enzimática (MCGINN, 2013).

Em 2005, o primeiro sequenciador de segunda geração foi apresentado, surpreendendo devido a maior quantidade de dados gerados e o custo drasticamente reduzido por base, podendo produzir um sequenciamento completo do genoma humano com cobertura de 309x em até 10 dias, com custo abaixo de R\$ 10.000,00 reais (GUT, 2013).

O foco central do sequenciamento de segunda geração é o desenvolvimento de sistemas que são muito baratos e muito rápidos com objetivo de gerar grande quantidades de dados para o sequenciamento de genomas, produzindo fragmentos curtos de 70 a 400 bases que serão reunidos computacionalmente, podendo gerar cerca de 60 Gb em um único dia de sequenciamento (GUT, 2013).

A terceira geração é definida pela leitura direta de moléculas de ácido nucleico individuais, no qual nenhum sistema enzimático de replicação é usado para identificar a sequência (GUT, 2013). Esse sistema consiste em um nanoporo incorporado em uma membrana (biológica ou sintética) imerso em solução salina. Então, uma corrente é aplicada ao sistema que transmite íons através do nanoporo, com a chegada de uma biomolécula carregada, por exemplo uma molécula de DNA, criará uma resistência no fluxo de íons que levará a mudanças na corrente elétrica que pode ser

medida (MCGINN, 2013). Moléculas de tamanho diferente bloqueiam o nanoporo por um grau diferente e por um tempo diferente, permitindo a sua identificação (BAYLEY, 2006).

Já a quarta geração é o método mais recente e sua aplicação é mais para a investigação de sequências de DNA já definidas, como as que possam ter sofrido mutação somática, ao invés da geração de sequências completas de cada célula (MCGINN, 2013).

Atualmente, o sequenciamento de DNA atingiu um ponto em que pode ser aplicado de forma holística (MCGINN, 2013). Nos últimos anos, ocorreu o desenvolvimento de diferentes técnicas de sequenciamento como os segunda, terceira e quarta geração; no entanto, uma nova geração não inutiliza a anterior, inclusive existe uma sobreposição desses diferentes métodos que ajuda a completar o conjunto de ferramentas necessário para uma análise genômica mais abrangente (GUT, 2013).

Apesar de todos esses métodos, o sequenciamento *de novo* de um genoma completo está longe de se tornar uma rotina, ainda são necessárias novas tecnologias para facilitar o sequenciamento de regiões genômicas repetitivas (incluindo transposons, sequências satélite, duplicações segmentares, sequências ribossômicas e outras sequências similares), que em grande parte, ainda “confundem” o sequenciador, dificultando o sequenciamento mesmo de genomas relativamente pequenos (SHENDURE, 2012).

Outro desafio para o futuro é o desenvolvimento técnicas que permitam mapas mais completos da variação genética, por exemplo, métodos atuais têm sensibilidade e especificidade limitadas para a detecção de pequenas inserções e deleções, expansões de repetições em *tandem*, eventos de transposição, variação do número de cópias, rearranjos neutros de cópia e todos os tipos de variação em regiões estruturalmente complexas de um genoma (SHENDURE, 2012).

Também devemos notar que, apesar de o advento do NGS ter revolucionado a pesquisa de genômica trazendo o sequenciamento de genomas inteiros ao alcance de muitos pequenos laboratórios, grande parte do alto rendimento do NGS deve-se aos avanços nos tratamentos dos dados gerados. Deste modo, as tecnologias NGS necessitam preponderantemente está associada a ferramentas computacionais, tanto para as análises quanto para o gerenciamento de dados (VAN DIJK, 2014).

3.4 SEQUENCIAMENTO DOS GENOMAS E SUAS APLICAÇÕES EM ESTUDOS EVOLUTIVOS

O sequenciamento de um genoma fornece um grande volume de dados genéticos, os quais podem ser utilizados para o estudo comparativo da organização genômica entre diferentes espécies (BOURQUE, 2005). Desta forma, revelou-se como uma excelente ferramenta para a descoberta de novos genes e suas respectivas funções (KIM, 2007), bem como para identificação de segmentos sintênicos e possíveis rearranjos cromossômicos envolvidos na organização genômica ao longo de uma escala evolutiva (BOURQUE, 2005).

Assim, mediante a análise *in silico*, esses dados facilitam a construção de mapas genéticos de alta densidade, possibilitando a identificação da sintonia e colinearidade, podendo ser utilizados em estudos filogenéticos (DIXIT, 2014). Assim, atualmente, o sequenciamento completo de um genoma é visto como uma das melhores ferramentas para a investigação de relações filogenéticas; no entanto, ainda existe um número limitado de organismos totalmente sequenciados (CAPOZZI, 2016).

Em termos gerais, o estudo de mapeamento comparativo *in silico* dos genes evidencia os segmentos ortólogos, ou seja, segmentos homólogos encontrados em diferentes espécies, também conhecidos como blocos ou segmentos sintênicos. Ou seja, os segmentos ou genes sintênicos são regiões que permanecem conservadas em cromossomos correspondentes de diferentes espécies (TANG, 2008), os quais podem ser utilizados como peças de um quebra cabeça para a reconstrução de genomas ancestrais (KEMKEMER et al., 2009), permitindo, a partir da definição dos rearranjos ocorridos, a busca de um cenário mais parcimonioso sobre a organização atual dos genomas dos organismos (BOURQUE, 2005). Tais rearranjos estão diretamente relacionados a quebra de colinearidade, a qual pode ser definida como a conservação dos genes na mesma ordem em um bloco sintênico (XU, 2016).

Uma das estratégias para a análise evolutiva de um genoma seria concentrar as análises apenas nas regiões gênicas ao invés de alinhar todo o genoma sequenciado, esse método reduz a complexidade do alinhamento do genoma inteiro, o que facilita a identificação dos segmentos sintênicos conservados (KEMKEMER et al., 2009). Os estudos de sintonia podem ser divididos em macrossintênicos e microssintênicos; a macrossintonia é uma forma de sintonia que envolve uma grande quantidade de genes em uma região na extensão do cromossomo, enquanto que a

microsintenia envolve apenas curtos segmentos do cromossomo com um número limitado de genes entre diferentes espécies (GAO, 2017).

Diferentes trabalhos mostraram que os blocos sintênicos não cobrem todo o cromossomo; normalmente, observamos a existência de ponto de interrupção entre eles, o qual é chamado de ponto de quebra (BOURQUE, 2004). Com isso, o mapeamento desses pontos de quebra permite avaliar se essas regiões coincidem ou não com possíveis pontos críticos de rearranjo cromossômico, isto é, sítios frágeis que podem estar associados a evolução (KEMKEMER et al., 2009).

Um estudo, utilizando alguns representantes de mamíferos, revelou que essas rupturas não ocorrem aleatoriamente; tal resultado indica a existência de regiões propensas a instabilidade nos genomas de mamíferos, as quais podem resultar em anormalidades cromossômicas ou se comportar como pontos propícios aos rearranjos cromossômicos, conduzindo aos eventos de especiação (KEMKEMER et al., 2009; SCHIBLER, 1998).

Alguns pontos de quebra de reutilização são regiões que, após a ruptura, tendem a formar um novo centrômero ou telômero para que o ciclo celular ocorra normalmente; sendo assim, foi revelado que os telômeros e os centrômeros tendem a se localizar em pontos de quebra, os quais são compartilhados durante a evolução; este mecanismo sugere que a evolução cromossômica de certos grupos é dirigida por quebras evolutivas, as quais ocorrem preferencialmente em regiões centroméricas e teloméricas (MURPHY, 2005).

Esses rearranjos cromossômicos têm sido identificados pela técnica citogenética comparativa de pintura cromossômica, a qual tem revelado a presença de segmentos conservados (macrossintenia) entre os cromossomos de primatas; entretanto, esta técnica apresenta uma limitação na inferência sobre a ordem dos genes dentro dos segmentos sintênicos (KEMKEMER et al., 2009).

Nas duas últimas décadas, os estudos citogenéticos e genômicos comparativos têm se tornado um campo de pesquisa altamente dinâmico e complementar, principalmente, devido aos avanços tecnológicos e a visão de que muitos dos conhecimentos sobre as funções do genoma podem ser melhores entendidas quando as informações evolutivas são levadas em consideração (MÜLLER, 2006).

Nesse contexto, a partir do sequenciamento e estudo comparativo de genomas, incluindo a pintura cromossômica, foi-se possível a construção de mapas genéticos de diferentes espécies de vertebrados; por exemplo, o sequenciamento em

larga escala do genoma do ser humano e do rato forneceu informações importantes sobre as taxas e os padrões da evolução cromossômica, permitindo a elucidação dos rearranjos que originaram os genomas de alguns mamíferos existentes (MURPHY; STANYON; O'BRIEN, 2001).

3.5 SEQUENCIAMENTO GENÔMICO DE *Callithrix jacchus*

Dentre os primatas, muitos questionamentos têm sido realizados a respeito das relações evolutivas dos macacos do Novo Mundo; neste grupo, apenas duas espécies apresentam projetos de genoma concluídos, o do macaco-esquilo (*Saimiri boliviensis*) e o do sagui-comum (*Callithrix jacchus*) (CAPOZZI, 2016).

O *C. jacchus* foi o primeiro macaco do Novo Mundo a ser sequenciado. Um genoma de 2,26 Gb de uma fêmea foi sequenciado, pelo método de *Whole-genome Shotgun*, representando cerca de 90% do genoma deste primata. Este genoma tem fornecido um poder estatístico sem precedentes para identificar sequências contrastantes entre primatas, facilitando a descoberta de regiões genômicas subjacentes à evolução fenotípica deste grupo de animais (*The Marmoset Genome Sequencing and Analysis Consortium*, 2014).

Características únicas, tais como modificações no microRNA expresso nas placentas e mutação não sinônimas em genes codificadores de proteínas envolvidos na fisiologia reprodutiva, foram notadas no genoma de *C. jacchus*, as quais podem estar relacionadas a frequência de gestações múltiplas e quimerismo. Em um total de 21.168 genes (44.973 transcritos), o número de 219 genes apresentou transcritos com suporte proteico para o sagui, enquanto que 15.706 genes não mostram evidência proteica para esta espécie, ou seja, sem proteínas descritas para esses genes neste organismo, mas sim, para humanos. Os 5.243 genes restantes apresentaram transcritos com suporte para dados proteicos a partir de outros organismos (*The Marmoset Genome Sequencing and Analysis Consortium*, 2014).

Tais dados do sequenciamento do genoma de *C. jacchus* encontram-se disponíveis em bancos de dados, além de terem sido montados em diferentes pseudocromossomos, os quais podem ser utilizado para guiar a construção de novos genomas e em análises comparativas sobre sua evolução utilizando métodos e modelos de bioinformática.

3.6 IMPORTÂNCIA DA BIOINFORMÁTICA PARA ANÁLISE DE DADOS BIOLÓGICOS

Nos últimos anos, o desenvolvimento rápido da genômica e proteômica gerou uma grande quantidade de dados biológicos e, para tirarmos conclusões a partir desses dados, necessitamos de análises computacionais sofisticadas; portanto, neste contexto, a bioinformática, ou biologia computacional, é a ciência interdisciplinar de interpretação de dados biológicos que utiliza a tecnologia da informação e ciência da computação (RAZA, 2012), incluindo conhecimentos de biologia molecular, estatística e matemática (RAUT, 2010).

O termo bioinformática foi criado por Paulien Hogeweg em 1979 para o estudo de processos informáticos em sistemas bióticos, sendo utilizado pela primeira vez na genômica e genética no final da década de 1980, principalmente nas áreas da genômica que envolviam o sequenciamento de DNA em grande escala (RAZA, 2012).

A bioinformática pode ser definida como a aplicação da tecnologia de computação no gerenciamento de informações biológicas, armazenando, extraíndo, organizando, analisando, interpretando e utilizando as sequências, especialmente de DNA, tendo principal objetivo aumentar a compreensão dos processos biológicos (RAZA, 2012) para que possamos modelar as macromoléculas, identificar os problemas de várias doenças e suas causas, projetar novos medicamentos e proteínas para melhorar a vida humana (RAUT, 2010).

Este campo de atuação começou apenas como uma ferramenta para gerenciar e analisar dados experimentais (ISEA, 2015); entretanto, com o surgimento de novas tecnologias e a grande quantidade de dados, a importância da bioinformática só tende a crescer possibilitando a aplicação e desenvolvimento de novas técnicas de mineração de dados (RAZA, 2012), por exemplo, permitindo um melhor mapeamento das variações interindividuais nas sequências de DNA, da expressão devido exposição ambiental e da susceptibilidade à doença, que são objetivos importantes para a genética humana e epidemiologia genética (MOORE, 2010).

Dentre as técnicas de mineração de dados para bioinformática podemos citar a descoberta de genes, detecção de domínio de função de proteína, detecção de motivos funcionais, inferência de função de proteína, diagnóstico de doenças,

prognósticos de doenças, otimização de tratamento de doenças e construção de redes de interações de proteínas e genes (RAZA, 2012).

Três componentes principais constituem a bioinformática: o primeiro, são os grandes bancos de dados, os quais permitem o armazenamento e gerenciamento de dados biológicos, como o NCBI (*National Center for Biotechnology Information*; Centro Nacional de Informações Biotecnológicas) (<https://www.ncbi.nlm.nih.gov>) e o *Ensembl* (<https://www.ensembl.org/index.html>); o segundo, são algoritmos e estatísticas, para determinar as relações entre esses grandes conjuntos de banco de dados; e o terceiro, são as ferramentas computacionais necessárias para analisar e interpretar os dados biológicos (RAUT, 2010), por exemplo, o MEGA 6 (TAMURA et al., 2013) e o Clustal Ômega (SIEVERS et al., 2011), ferramentas utilizadas em alinhamentos globais e construção de dendogramas.

Desde a sua criação, a bioinformática vem mudando o cenário da pesquisa e oferece oportunidades para que os cientistas utilizem metodologias melhoradas, aprimorem o uso de dados e abordem rapidamente questões de pesquisa importantes (FORMAN, 2010), como o desenvolvimento de técnicas de detecção e terapias de câncer causados por alterações genéticas e anormalidades moleculares, como mutações, translocações e deleções (LU, 2015).

4. MATERIAL E MÉTODOS

4.1 MINERAÇÃO DE DADOS PARA AS ANÁLISES DE MACROSSINTENIA

A partir de uma pesquisa realizada no *Genome Databases* [NCBI (*National Center for Biotechnology Information*; Centro Nacional para Informação Biotecnológica – <http://www.ncbi.nlm.nih.gov>], no dia 30 de outubro de 2017 e usando a palavra-chave *Callithrix jacchus jacchus*, obteve-se uma RefSeq NC_013919.1 para o cromossomo Y de *Callithrix jacchus*, a partir do genoma de referência *Callithrix jacchus Callithrix jacchus-3.2* (<https://www.ncbi.nlm.nih.gov/genome>). Em seguida, utilizando a citada RefSeq, foram catalogados os genes presentes no cromossomo sexual Y de *Callithrix jacchus*, sendo selecionados dentre eles, aqueles genes que possuíam vínculo ao *Ensembl Release 90* (<http://www.ensembl.org>), disponível no *Genome Data Viewer*, anotando-se a identificação, tamanho em número de bases, função e sua localização (APÊNDICE A, Quadro 1).

Em seguida, após a caracterização do cromossomo sexual de *Callithrix jacchus*, foram selecionadas as espécies entre as quais esse cromossomo seria comparado. Os critérios utilizados foram: as espécies deveriam apresentar dados citogenéticos dos autossomos e, pelo menos, um dos cromossomos sexuais publicados e relacionados ao *Ensembl Release 90* disponíveis no *Genome Data Viewer*, com o intuito de se observar possíveis relações sintênicas entre autossomos e cromossomos sexuais. Dos 39 genomas descritos no NCBI, até o dia de acesso (03 de novembro de 2017), apenas cinco espécies de primatas se encaixaram nessas condições, o *C. jacchus* (sagui-comum; macaco do Novo Mundo), *Chlorocebus sabaeus* (Linnaeus, 1766) (macaco-verde; macaco do Velho Mundo), *Homo sapiens* (Linnaeus, 1758) (ser humano; Hominidae), *Macaca mulatta* (Zimmermann, 1786) (macaco *Rhesus*; macaco do Velho Mundo) e *Pongo abelii* (Lesson, 1827) (Orangotango-de-sumatra; Hominidae). Além destes, um representante da ordem Artiodactyla, *Sus scrofa* (Linnaeus, 1758) (Javali; Suidae), foi escolhido como grupo externo nestas análises.

Posteriormente, utilizando o genoma *Marmoset (C_jacchus3.2.1)* no *Ensembl Release 90*, os genes ortólogos de cada gene do cromossomo Y de *C. jacchus* referente as cinco espécies escolhidas foram catalogados, observando a identificação, tamanho em número de bases, função e sua localização para as análises de

macrossintênicas. Os dados minerados foram listados em quadros e a ordem dos genes foi sempre respeitada, ou seja, o gene número 10 de uma espécie sempre equivale ao ortólogo do gene número 10 de outra espécie (APÊNDICE A, Quadros 2, 3, 4, 5 e 6).

4.2 CONSTRUÇÃO DOS MAPAS GENÉTICOS E ANÁLISES DE MACROSSINTÊNICAS

Para a construção dos mapas genéticos dos cromossomos utilizou-se a ferramenta *online* MapGene2Chrom web v2 (http://mg2c.iask.in/mg2c_v2.0/) (JIANGTAO, 2015), levando em consideração os dados de localização e identificação dos genes ortólogos entre as espécies; essa identificação foi realizada mediante a sigla do gene dada pelos bancos de dados acessados ou pelo número no catálogo determinado no presente trabalho. Este último número foi determinado, uma vez que alguns genes anotados estavam nomeados de formas diferentes entre as espécies; por este motivo, foi padronizado um código para o gene e seus ortólogos. Por exemplo, o gene 2 equivale ao gene ENSCJAG000000130 56 em *C. jacchus* e seus ortólogos ENSG00000178605 em *H. sapiens*, ENSPPYG00000020075 em *P. abelli* e ENSSSCG00000033858 em *S. scrofa* (Quadro 7).

A partir dos mapas genéticos, foram observados os segmentos sintênicos e, conseqüentemente, foram levantadas hipóteses para possíveis rearranjos envolvendo os cromossomos em que estes segmentos foram localizados, a exemplo de inversões e translocações. Vale salientar que a ordem e composição gênica dos cromossomos foi considerada para a determinação do rearranjo ocorrido.

5. RESULTADOS E DISCUSSÃO

5.1 CARACTERIZAÇÃO DO CROMOSSOMO Y

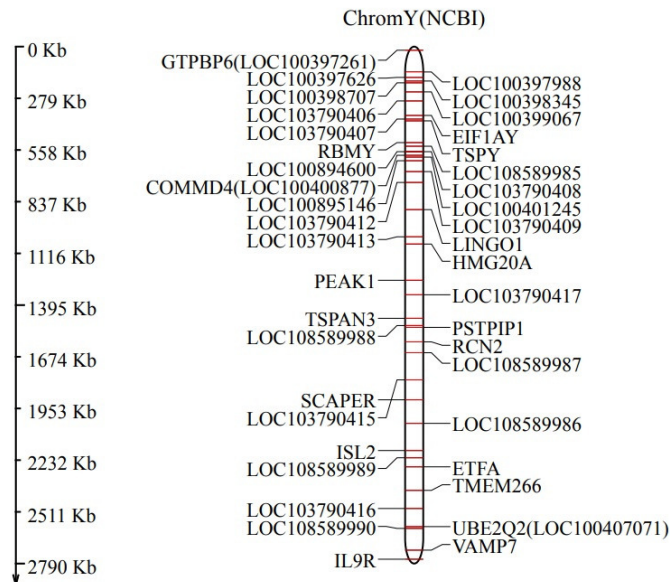
O RefSeq NC_013919.1 do cromossomo Y de *Callithrix jacchus* apresenta 2.853.901 bases de extensão, com 41 genes descritos (Tabela 1; Figura 2). Em comparação as outras espécies analisadas, nesse cromossomo, puderam ser observadas diferenças quanto ao tamanho e quantidade de genes. De acordo com o NCBI, o cromossomo Y possui 57.227.415 de bases de extensão e 577 genes em *Homo sapiens* (Hsap) (NC_000024.10), tem 11.753.682 de bases de extensão e 92 genes em *Macaca mulatta* (Mmul) (NC_027914.1), tem 6.181.219 de bases de extensão e 51 genes em *Chlorocebus sabaeus* (Csab) (NC_023672.1), ainda não foi descrito em *Pongo abelii* (Pabe), enquanto que para o *Sus scrofa* (Sscr) (NC_010462.3) foi verificado um cromossomo de 43.547.828 bases de extensão e 147 genes (Tabela 1).

Tabela 1 – Espécies analisadas, identificação das sequências de referência (RefSeq) para os cromossomos Y destas espécies, bem como a sua extensão em bases e quantidade de genes anotados, segundo o NCBI (acesso em 30 de outubro de 2017).

Espécie	RefSeq	Extensão em bases	Quantidade de genes
<i>Callithrix jacchus</i>	NC_013919.1	2.853.901	41
<i>Homo sapiens</i>	NC_000024.10	57.227.415	577
<i>Macaca mulatta</i>	NC_027914.1	11.753.682	92
<i>Chlorocebus sabaeus</i>	NC_023672.1	6.181.219	51
<i>Sus scrofa</i>	NC_010462.3	43.547.828	147

Os cromossomos X e Y em primatas divergiram de um par de autossomos ancestral e, ao longo do tempo pela ausência de recombinação meiótica, o cromossomo Y acumulou uma série de deleções mutacionais tornando-o estratificado, vulnerável a perda gênica e a degeneração evolutiva em comparação ao cromossomo X (HUGHES; ROZEN, 2012; MARSH, 2014). O cromossomo Y tem evoluído de forma independente em diferentes grupos de plantas e animais, mostrando-se bastante diverso (HUGHES; ROZEN, 2012), como notado para as espécies aqui analisadas.

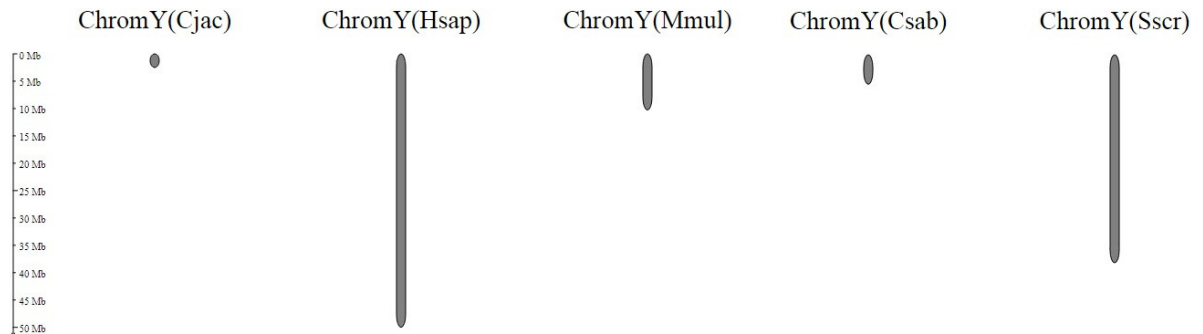
Figura 2 - Mapa genético do cromossomo Y de *C. jacchus* contendo 41 genes e construído a partir de informações disponíveis no NCBI, acesso em 30 de outubro de 2017, mediante a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.



Comparando-se os dados, em termos de extensão e quantidade de genes, nota-se que *Cjac* apresenta o menor cromossomo Y e menor quantidade de genes descrita, diferenciando dos Catarrhini, cujo representante *Hsap* mostrou o maior cromossomo Y e a maior quantidade de genes (Figura 3), fato que pode ser explicado pela presença de um grande bloco heterocromático terminal derivado neste cromossomo em humanos (STANYON et al., 2008). Entretanto, tais divergências em termos de extensão e composição gênica podem ser discutidas sobre o ponto de vista biológico, o qual está associado a natureza complexa e repetitiva dos cromossomos Y, bem como tecnológico, levando-se em consideração os métodos de sequenciamento gênico (MARSH, 2014).

Este cromossomo sexual tem sido dividido em duas regiões principais: as Pseudoautossômicas (PARs) e Macho-específica (MSY) (MARSH, 2014). As PARs são regiões terminais homólogas ao cromossomo X, presentes nos braços curto e longo, sendo assim as únicas regiões do cromossomo Y passíveis de recombinação meiótica (MARSH, 2014; RAPPOLD, 1993). Por sua vez, a região MSY codifica todo o material genético que é responsável pela diferenciação sexual masculina, o qual é transmitido clonalmente, de pai para filho (HUGHES, 2012), não sofrendo recombinação genética.

Figura 3 - Comparação do tamanho em bases de extensão dos cromossomos Y das espécies em estudo, *Callithrix jacchus* (Cjac), *Homo sapiens* (Hsap), *Macaca mulatta* (Mmul), *Chlorocebus sabaeus* (Csab) e *Sus scrofa* (Sscr), os quais foram construídos mediante a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.



No presente trabalho, tais regiões foram identificadas pela presença de genes específicos, a exemplo de *GTPBP6* e *IL9R* localizados na extremidade superior e inferior, respectivamente, do mapa genético construído para *C. jacchus*, sendo estes apontados na literatura como genes situados nas PARs (CHARCHAR, 2003; VAWTER, 2007; VERMEESCH, 1997). Por sua vez, a região MSY foi identificada pela presença dos genes específicos do cromossomo Y, a exemplo de *RBMV* e *TSPY* (ELLIOTT, 1997; LAU, 1999).

Estudos demonstraram que a região MSY dos Eutérios é subdividida em zonas distintas: a região heterocromática, área constituída de sequências repetitivas cujo DNA está altamente compactado e atua na regulação da expressão gênica; região X-transportada; região X-degenerada; e, região amplicônica. As duas últimas regiões são os principais componentes eucromáticos do cromossomo Y, possuindo repertórios gênicos e trajetórias evolutivas distintas (LI, 2013; MARSH, 2014).

As sequências X-degeneradas são uma versão deteriorada do cromossomo X, sendo compostas por genes cópia única, por exemplo 16 genes em humanos, as quais são expressas ao longo do corpo e provavelmente desempenham funções *housekeeping* celular básicas (BACHTROG, 2013; MARSH, 2014). Por sua vez, as regiões amplicônicas são altamente repetitivas, compostas por longos trechos de sequências duplicadas, as quais estão organizadas em estruturas palindrômicas que albergam famílias gênicas multicópias com um padrão de expressão específico de testículos (HUDHES; ROZEN, 2012).

As funções desempenhadas pelo conteúdo gênico de cada uma dessas classes de sequências fazem com que elas sejam submetidas a diferentes pressões seletivas. Assim, em termos de conservação, as sequências X-degeneradas possuem uma tendência a serem mais conservadas entre primatas, enquanto que as sequências amplicônicas têm acumulado mais mutações e evoluído mais rapidamente (MARSH, 2014), a exemplo do notado ao se comparar humanos, chimpanzé e macaco Rhesus. Nestas espécies, as sequências amplicônicas sofreram uma drástica reestruturação e renovação, seguindo trajetórias evolutivas completamente diferentes em cada linhagem (HUGHES; ROZEN, 2012).

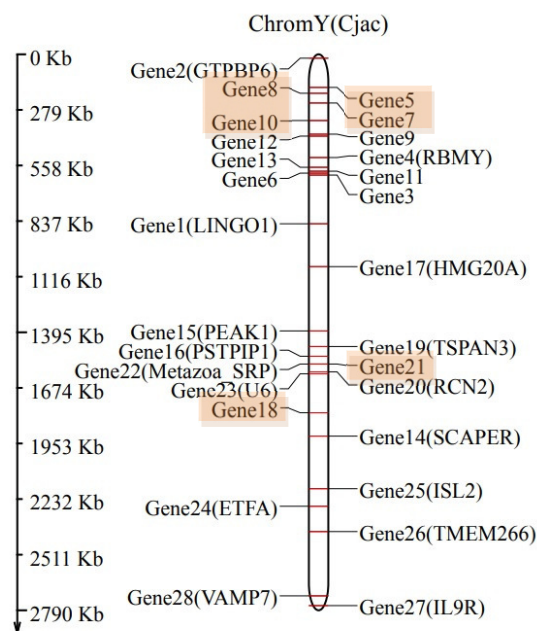
A ausência de recombinação meiótica na região MSY interfere diretamente na atuação da seleção natural. É fato que o coeficiente de seleção e as taxas de mutação para alelos benéficos e/ou deletérios influenciam na perda de genes nos cromossomos Y, o que tem levado a diminuição deste cromossomo ao longo do processo evolutivo (BACHTROG, 2013; HUGHES, 2015). Entretanto, a sobrevivência deste cromossomo tem sido relacionada com a ocorrência de conversão gênica nas regiões amplicônicas; estas, compostas por famílias multigênicas, fornecem alvos para recombinações inter e intracromátides gerando inversões, deleções e duplicações, as quais podem proteger genes importantes de efeitos deletérios e, ao mesmo tempo, contribuir para a existência de diferentes classes de sequências MSY entre as espécies (variabilidade genética) (HUGHES; ROZEN, 2012; HUGHES, 2015; MARSH, 2014).

Neste sentido, sugere-se que ausência de recombinação meiótica, a seleção natural e mecanismo de conversão gênica podem estar relacionados as diferenças no tamanho cromossômico e na quantidade de genes presentes no cromossomo Y das espécies avaliadas (Figura 3). Já no contexto tecnológico, devido a natureza repetitiva deste cromossomo Y, talvez a diferença da quantidade de genes entre as espécies analisadas esteja relacionada a dificuldade de reunir fragmentos de sequências repetitivas, inclusive as amplicônicas, resultantes do sequenciamento pelo método WGS e, conseqüentemente, interferir na identificação de sequências gênicas (HUGHES, 2015; HUGHES; ROZEN, 2012).

5.2 ANÁLISE MACROSSINTÊNICA DO CROMOSSOMO Y DE *Callithrix jacchus* FRENTE OUTROS PRIMATAS

Dentre os 41 genes identificados no cromossomo Y de *C. jacchus* no NCBI, apenas 28 foram visualizados no *Ensembl Release 90* pelo *Genome Data Viewer* (Figura 4). Entretanto, seis genes não apresentaram relações ortólogas descritas até o momento (Quadro 7), os genes identificados como 5, 7, 8, 10, 18 e 21.

Figura 4 - Mapa genético do cromossomo Y de *C. jacchus* contendo 28 genes segundo o *Ensembl release 90*. Tal mapa foi construído utilizando-se a ferramenta *online* MapGene2Chrom. Em destaque, estão os genes sem ortologia. Fonte: próprio autor.



Utilizando os genes ortólogos identificados para o cromossomo Y mediante o *Ensembl release 90*, foi possível a construção de mapas genéticos das espécies *Callithrix jacchus* (Cjac) (Figura 4), *Homo sapiens* (Hsap) (Figura 5), *Pongo abelii* (Pabe) (Figura 6), *Macaca mulatta* (Mmul) (Figura 7), *Chlorocebus sabaeus* (Csab) (Figura 8) e *Sus scrofa* (Sscr) (Figura 9). A análise comparativa destes mapas genéticos permitiu a identificação de macrossintenia, estando os genes ortólogos distribuídos em diferentes autossomos e cromossomos sexuais em todas espécies em estudo.

Em Hsap, o número de 27 genes foi localizado nos cromossomos 2 (um gene), 15 (13 genes), X (três genes) e Y (11 genes pertencentes as famílias multigênicas *TSPY* e *RBMY*). Por sua vez, para Pabe, o número total de 17 genes foi encontrado ao longo dos cromossomos 2 (um gene), 15 (13 genes) e X (três genes), dados similares ao notado para Hsap, quanto a quantidade e conteúdo gênico para os cromossomos 2 (Gene 23) e X (Gene 2, 27 e 28). Com relação ao cromossomo 15, entre os 13 genes anotados, onze foram comuns a ambas as espécies (Genes 6, 11, 14, 15, 16, 17, 19, 20, 24, 25 e 26) (Figuras 5 e 6).

Figura 5 – Mapa genético dos cromossomos de *Homo sapiens* (Hsap) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release 90*. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.

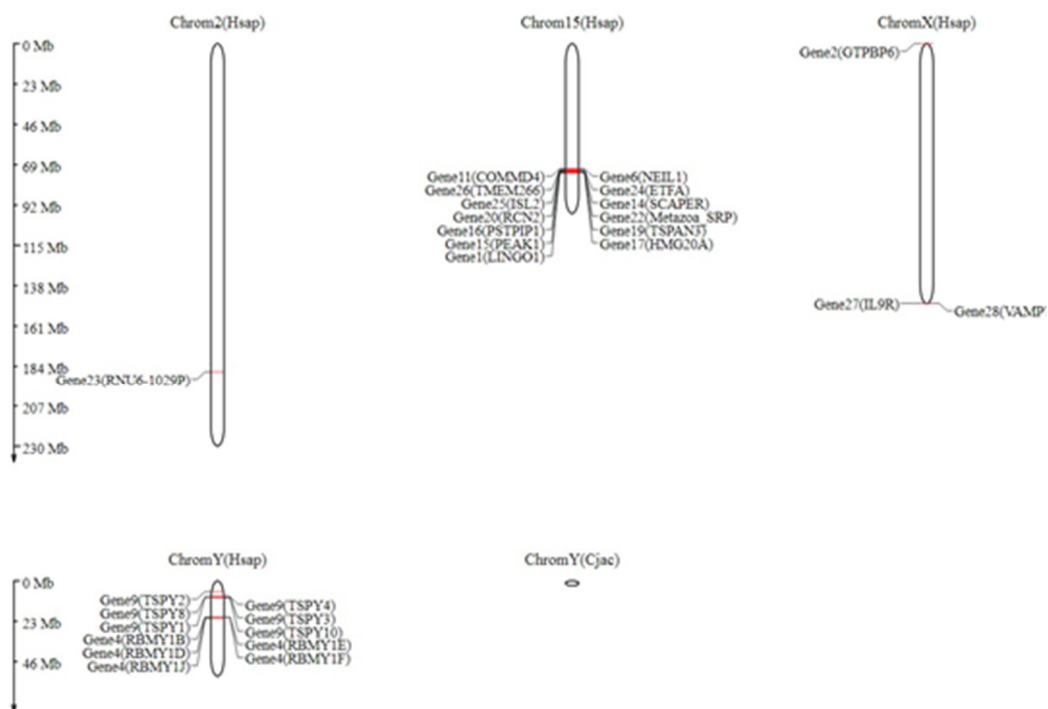
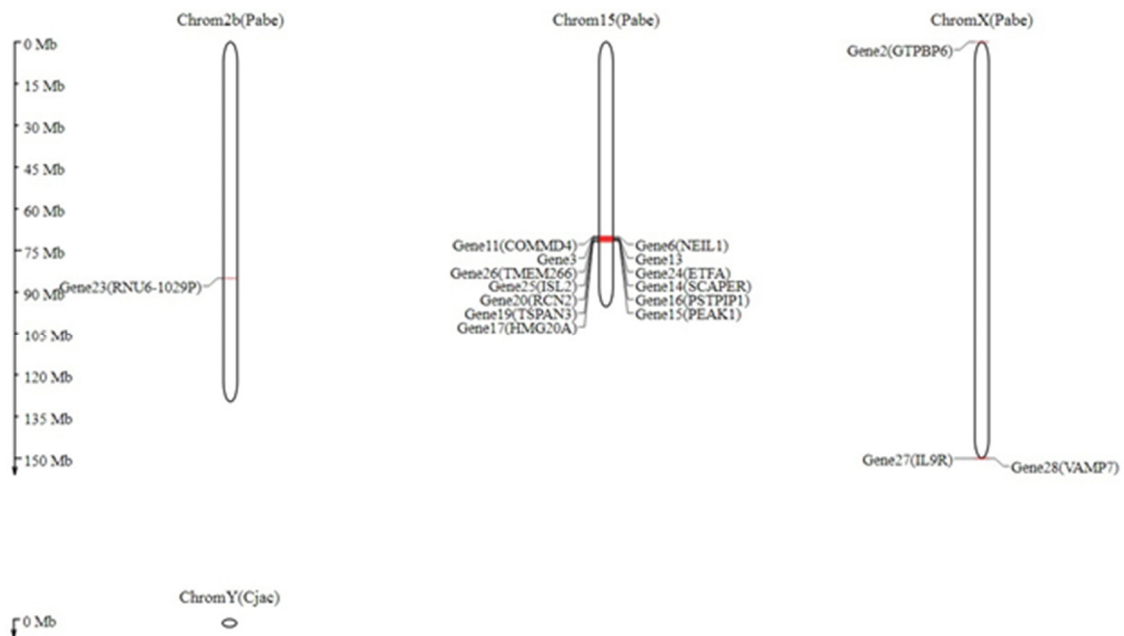


Figura 6 – Mapa genético dos cromossomos de *Pongo abelii* (Pabe) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release 90*. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.



Considerando o mapa de Mmul, vinte e dois genes foram identificados nos cromossomos 7 (13 genes), X (dois genes) e Y (sete genes pertencentes as famílias gênicas *TSPY* e *RBMY*), enquanto que em Csab, o número de 19 genes foi localizado entre os cromossomos 9 (um gene), 26 (14 genes), X (dois genes) e Y (dois genes, entretanto apenas um gene representante de *RBMY*). Vale ressaltar que os cromossomos 7 e 26 de Mmul e Csab, respectivamente, apresentam 12 genes em comum (1, 3, 6, 11, 14, 15, 16, 17, 19, 20, 22 e 24) (Figuras 7 e 8). Já em Sscr, dezenove genes foram mapeados nos cromossomos 1 (um gene), 3 (um gene), 7 (12 genes), 13 (um gene), X (um gene) e Y (três genes, sendo dois representantes de *TSPY* e *RBMY*) (Figura 9).

Comparando-se os representantes de Hominoidea e macacos do Velho Mundo, notou-se que o Hsap, Pabe, Mmul e Csab apresentam em comum os genes ortólogos 27 (*IL9R*) e 28 (*VAMP7*) no cromossomo X; entretanto, o gene 2 (*GTPBP6*), o qual foi mapeado no cromossomo X de Hsap e Pabe, não apresentou sequências ortológicas para as espécies Mmul e Csab (Figuras 5, 6, 7 e 8). Em relação ao cromossomo Y, foram mapeadas os genes ortólogos 4 (*RBMY*) e 9 (*TSPY*) para todas as espécies, exceto Csab que apresentou sequência ortóloga apenas para o *RBMY*. Adicionalmente, nove genes ortólogos (6, 11, 14, 15, 16, 17, 19, 20 e 24) ao

cromossomo Y de Cjac estavam localizados em cromossomos autossomos de todas as espécies analisadas.

Assim, levando em consideração a maior quantidade de genes ortólogos, o cromossomo Y de Cjac estava principalmente relacionado ao cromossomo 15 dos Hominoidea (Hspa e Pabe; Figuras 5 e 6), ao cromossomo 7 de Mmul (Cercopithecoidea; Figura 7) e Sscr (Cetartiodactyla; Figura 9) e ao cromossomo 26 de Csab (Cercopithecoidea; Figura 8).

O perfil genético similar desses cromossomos (7, 15 e 26) pode ser um indício de uma grande relação evolutiva entre eles, indicando um conteúdo genético similar e que eventos de evolução cromossômica ocorreram nos grupos estudados, a exemplo de inversões e translocações, os quais já foram relatados dentro dos Hominoidea e entre outros primatas (BIGONI, 1997; FARRÉ, 2012; MÜLLER, 2006; STANYON, 2012).

Figura 7 – Mapa genético dos cromossomos de *Macaca mulatta* (Mmul) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release 90*. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.

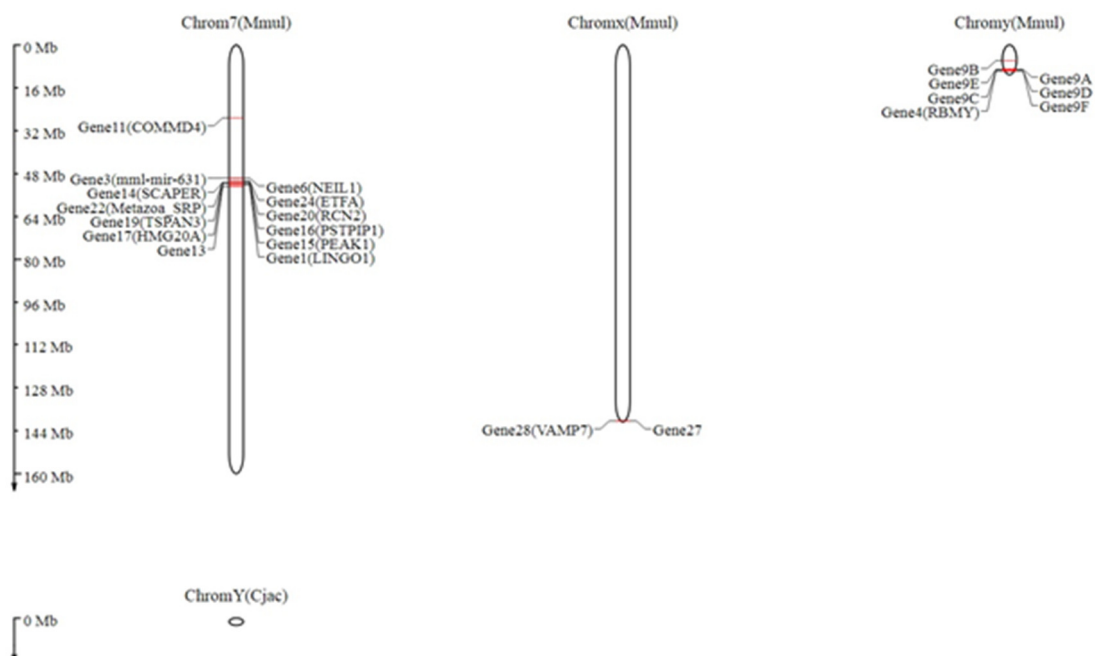


Figura 8 – Mapa genético dos cromossomos de *Chlorocebus sabaues* (Csab) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release 90*. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.

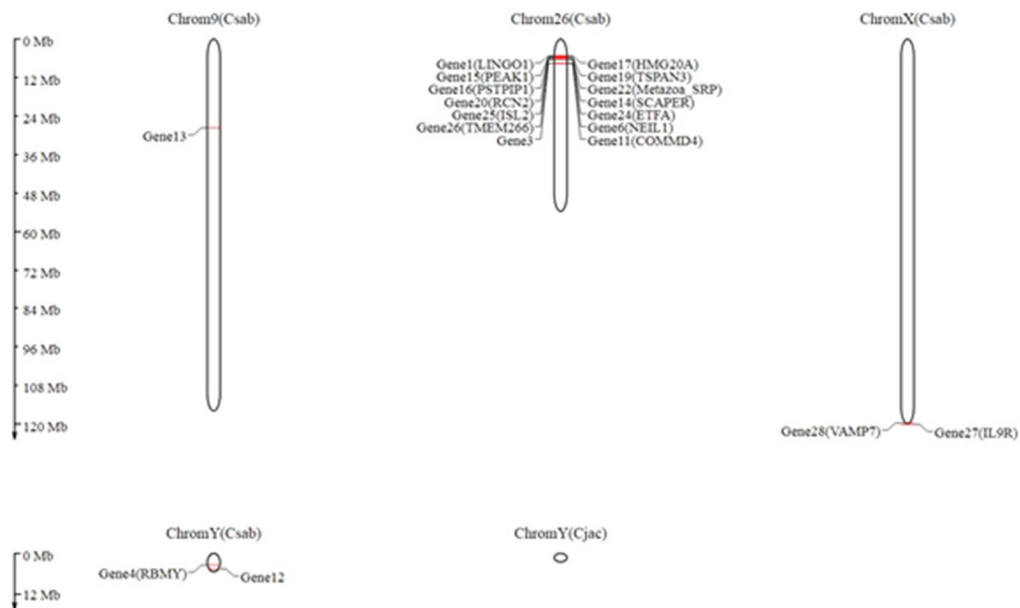
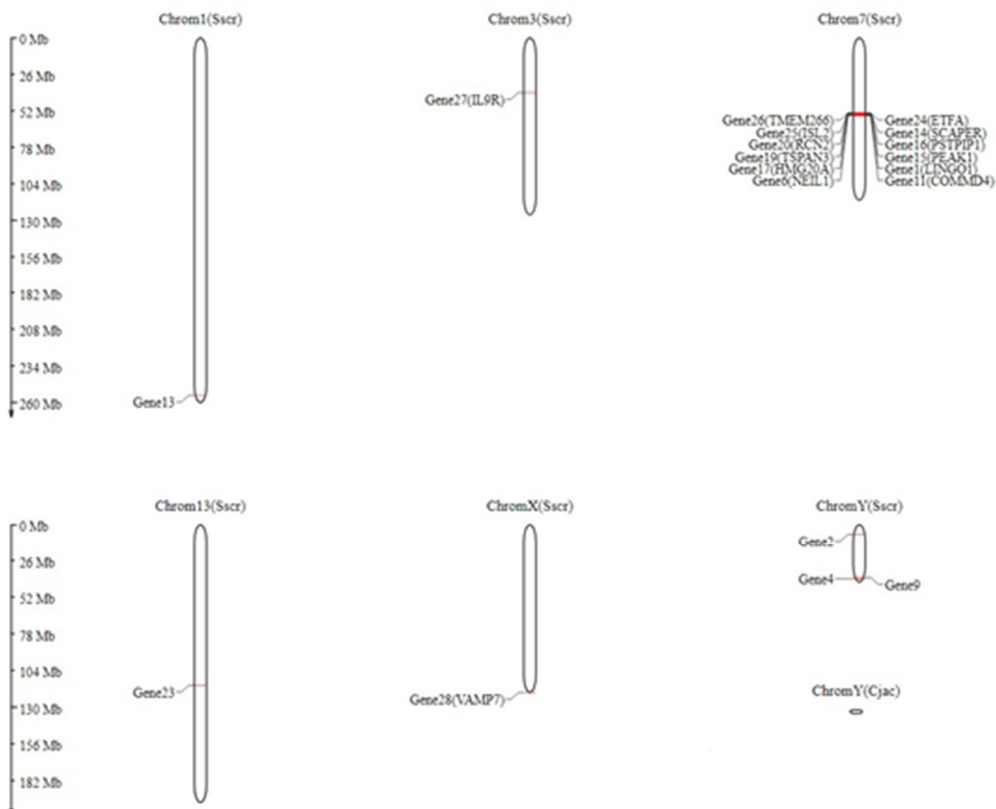
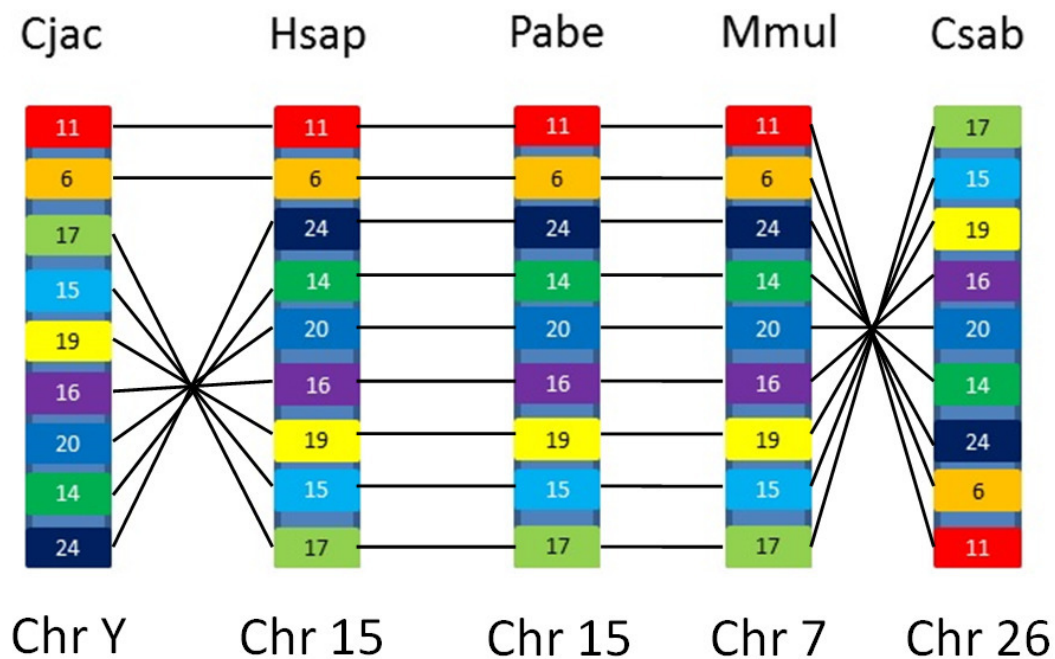


Figura 9 – Mapa genético dos cromossomos de *Sus scrofa* (Sscr) com genes ortólogos relacionados a *Callithrix jacchus* segundo o *Ensembl Release 90*. O cromossomo Y de *C. jacchus* [ChromY(Cjac)] foi representado apenas para a comparação do tamanho. Mapas construídos utilizando-se a ferramenta *online* MapGene2Chrom. Fonte: próprio autor.



A ordem dos genes órtologos distribuídos nos cromossomos Y em Cjac, 15 em Hsap e Pabe, 7 em Mmul e 26 em Csab foi identificada nos respectivos cromossomos, levando em consideração a sua localização, com o intuito de verificar a existência de colinearidade. Os resultados indicam quebras de colinearidade entre tais espécies, as quais podem ser decorrentes de rearranjos cromossômicos, a exemplo de translocações e inversões (Figura 10).

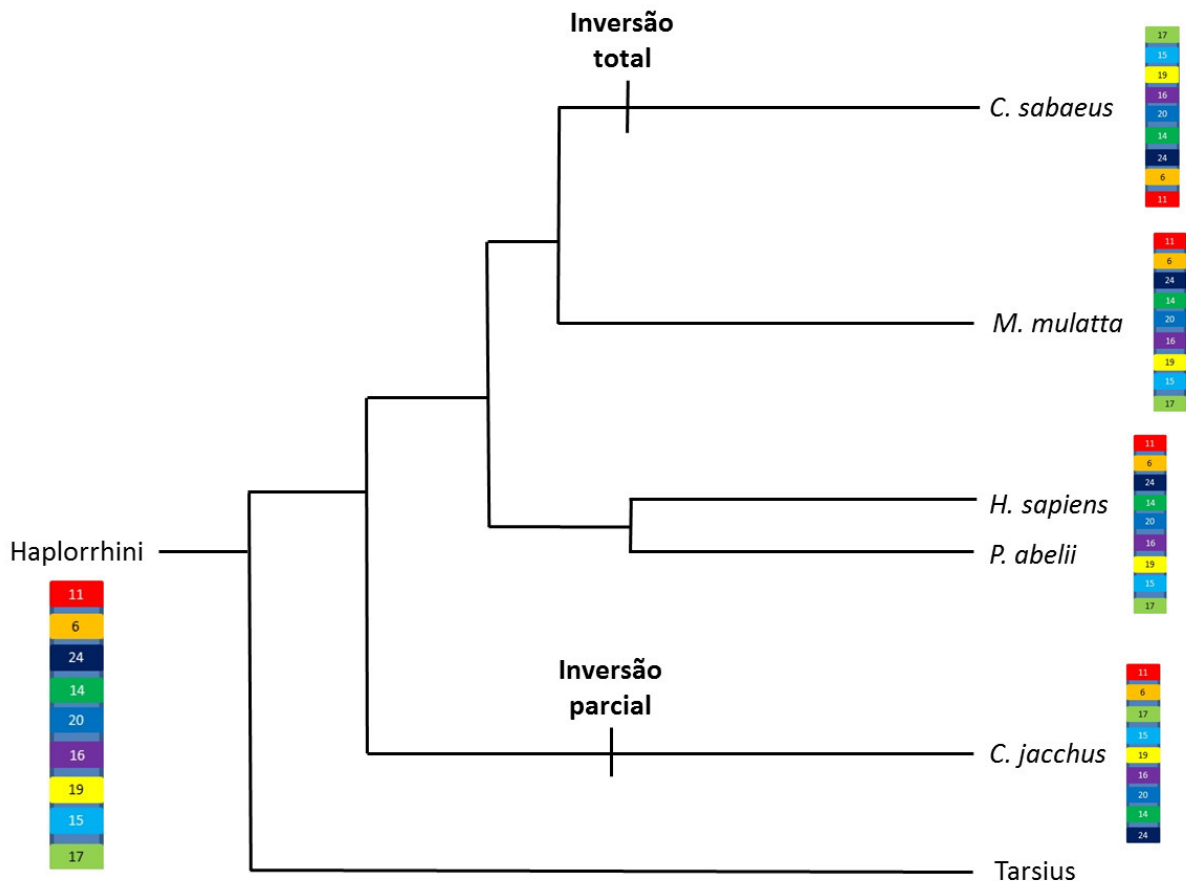
Figura 10 – Ordem dos genes ortólogos comuns as espécies *Callithrix jacchus* (Cjac), *Homo sapiens* (Hsap), *Pongo abelli* (Pabe), *Macaca mulatta* (Mmul) e *Chlorocebus sabaeus* (Csab), conforme as suas respectivas localizações disponíveis no *Ensembl Release 90*, mostrando quebra de colinearidade. Fonte: próprio autor.



Com base na filogenia descrita para os primatas (PERELMAN, 2011) e o princípio de parcimônia, sugere-se a existência de um cromossomo autossomo cuja ordem dos genes seria 11, 6, 24, 14, 20, 16, 19, 15 e 17 em um ancestral dos Haplorrhini (Figura 11). Após a divergência dos Platyrrhini deste ancestral, a qual ocorreu a 48,5 milhões de anos, teriam ocorrido rearranjos cromossômicos entre este autossomo e o cromossomo Y, a exemplo de translocação e inversão, notando-se a disposição atual dos genes em estudo no cromossomo Y de Cjac. Neste sentido, duas hipóteses podem ser levantadas: na primeira, inicialmente, teria ocorrido uma translocação desses genes do autossomo para o cromossomo Y, seguida de uma quebra e uma reunião invertida (inversão) (Figura 12); e na segunda, uma quebra e

reunião invertida teria ocorrido no autossomo e, posteriormente, uma translocação dessa região contendo os genes em questão para o cromossomo Y (Figura 13).

Figura 11 – Possíveis rearranjos cromossômicos ocorridos envolvendo os genes ortólogos comuns presentes nos cromossomos Y de *Callithrix jacchus* (Cjac), 15 de *Homo sapiens* (Hsap) e *Pongo abelli* (Pabe), 7 de *Macaca mulatta* (Mmul) e 26 de *Chlorocebus sabaesus* (Csab), levando-se em consideração a filogenia proposta por PERELMAN (2011). Fonte: próprio autor.



A respeito das translocações do cromossomo Y com os autossomos, de acordo com a literatura, esses eventos são raros entre os primatas. Entretanto, foi observado em *Presbytis cristata* (BIGONI, 1997), *Callimico goeldi* (DUMAS, 2017), *Homo sapiens* e *Pan troglodytes* (MÜLLER, 2006), tanto em *P. cristata* quanto em *H. sapiens* e *P. troglodytes*, esse tipo de translocação já foi relatado entre o cromossomo 1 e o cromossomo Y.

Figura 12 – Primeira representação hipotética de rearranjos cromossômicos entre um cromossomo autossomo ancestral e o cromossomo Y de *Callithrix jacchus*, com destaque para uma translocação seguida de inversão, resultando na ordem dos genes observadas neste cromossomo sexual. Fonte: próprio autor.

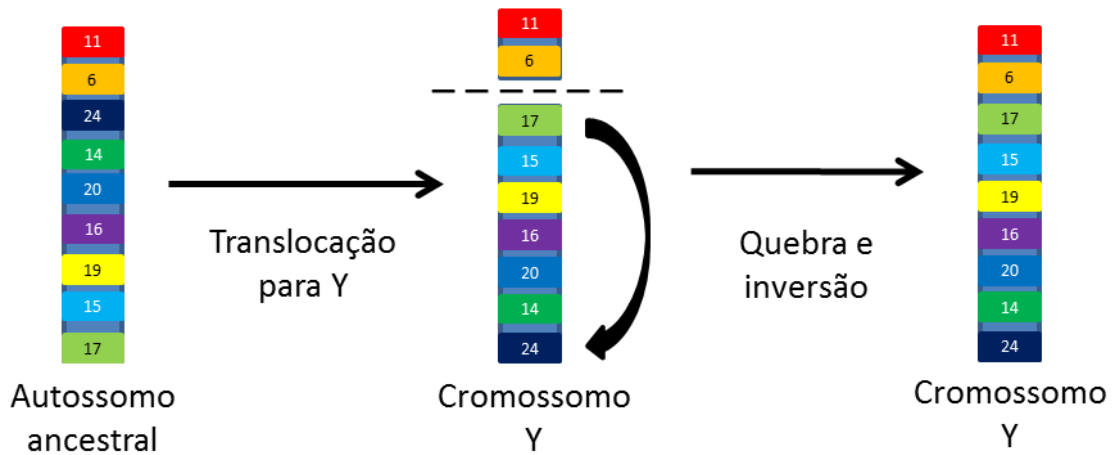
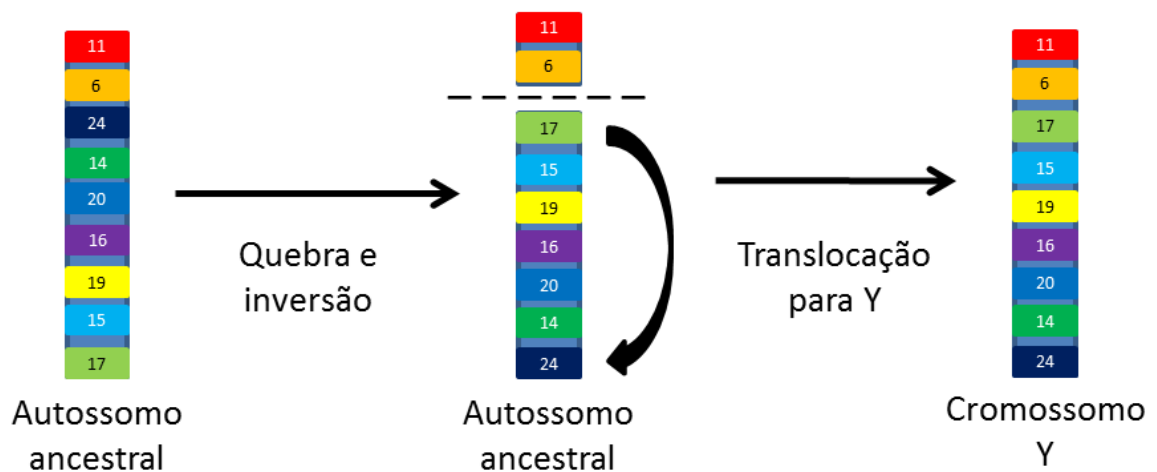
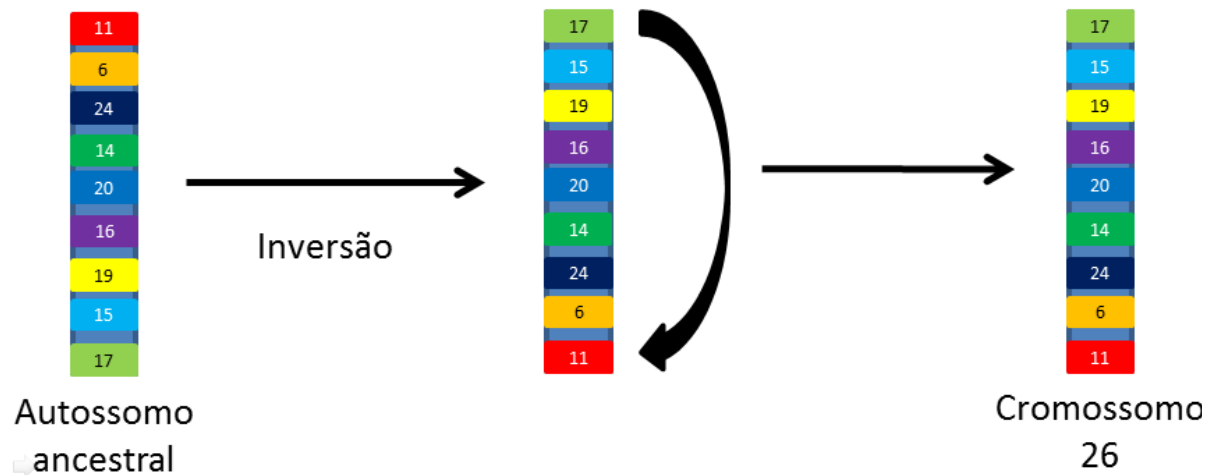


Figura 13 – Segunda representação hipotética de rearranjos cromossômicos entre um cromossomo autossomo ancestral e o cromossomo Y de *Callithrix jacchus*, com destaque para uma inversão no autossomo e, posterior, translocação para o cromossomo sexual, resultando na ordem dos genes observadas no cromossomo Y. Fonte: próprio autor.



Em relação aos Catarrhini, esta ordem foi mantida em Hsap e Pabe, ambos no cromossomo 15, bem como no cromossomo 7 de Mmul (Figura 11). De acordo com Stannyon et al (2008), estudos de macrossintenia com BACs (*Bacterial Artificial Chromosome*; Cromossomos Artificiais de Bactérias) revelaram a existência de sintenia entre os cromossomos 15 e 7 de Hsap e Mmul, respectivamente (NEUSSER, 2001). Entretanto, para Csab, observou-se que a ordem destes genes está inserida de forma invertida no cromossomo 26 (Figura 11), sugerindo que ao longo da evolução da linhagem de Csab, um inversão teria ocorrido (Figura 14).

Figura 14 – Representação hipotética de uma inversão no cromossomo autossomo 26 de *Chlorocebus sabaes*, enfatizando a quebra de colinearidade frente a ordem dos genes em questão visualizadas em cromossomos autossomos de *Homo sapiens*, *Pongo abelii* e *Macaca mulata*. Fonte: próprio autor.



Estudos de pintura cromossômica recíproca já identificaram associação de segmentos cromossômicos humanos a exemplo de 9/13, 9/22, 13/17 em espécies de *Callithrix*, comprovando a existência de rearranjos cromossômicos e macrossintenia entre esses primatas. Outra associação foi relatada entre os cromossomos 2/15 de humanos, exclusivamente, nos cariótipos de *Callithrix* e *Saimiri*. Tal informação enfatiza a ocorrência de rearranjos entre os cromossomos 2 e 15 em humanos, o que pode vir a refletir na presença de genes órtologos do cromossomo Y de Cjac nos referidos cromossomos de Hsap e Pabe (NEUSSER, 2001).

Os genes órtologos que foram localizados nos cromossomos X de Hsap, Pabe, Mmul e Csab (*GTPBP6* e *IL9R*) são genes referidos na literatura como presentes nas regiões pseudoautossômicas, região de homologia entre os cromossomos sexuais. Desta forma, pode ter ocorrido recombinação entre o

cromossomo X e Y de Cjac, motivo pelo qual estes genes estejam presentes no cromossomo Y desta espécie.

O gene *GTPBP6* está localizado na região pseudo-autossômica dos cromossomos X e Y, sendo associado a síndrome de Klinefelter e com cognição verbal (VAWTER, 2007). Segundo o *Ensembl Release 91*, este gene é um membro *GTP BINDING 6*, família de proteínas que se ligam ao GTP com importante função na sinalização de vias que regulam diferentes funções, como a fusão e fissão das vesículas de transporte (BALCH, 1990). Por sua vez, o gene *IL9R* (Interleukin-9 receptor) é um receptor para a citocina interleucina 9 (IL-9), expresso em muitas células hematopoéticas como as células T (DEMOULIN, 1998), localizado no braço longo do cromossomo X na maioria dos primatas, em sua região pseudoautossômica próximo ao telômero (CHARCHAR, 2003; VERMEESCH, 1997).

Para os genes presentes no cromossomo Y, foram os citados *TSPY* e *RMBY*, ambos localizados nas regiões macho-específica, mantendo a sintonia e colinearidade entre as espécies de primatas. O *TSPY* (*testis-specific Y-encoded*) é um gene evolutivamente conservado no cromossomo Y dos mamíferos (LAU, 1999), localizado principalmente em *locus* com predisposição para o desenvolvimento de câncer, conhecido como gonadoblastoma no cromossomo Y (GBY) (KIDO, 2006).

Por fim, o gene *RMBY* é encontrado no cromossomo Y dos mamíferos e relacionado a divisão meiótica das células germinativas masculinas, sua ausência já foi associada a não formação de gametas haploides (ELLIOTT, 1997). De acordo com o *Ensembl Release 91*, este gene faz parte da família proteica *Rna Binding Motif Protein Y Chromosome Family 1 Member*, cujos membros codificam uma proteína nuclear específica para as células germinativas masculinas (KUROIWA, 2010).

6. CONSIDERAÇÕES FINAIS

1. Foram observados e catalogados 41 genes para o cromossomo Y de *Callithrix jacchus* (RefSeq NC_013919.1), de acordo com o NCBI (*National Center for Biotechnology Information*; Centro Nacional para Informação Biotecnológica), entre os quais 28 também foram encontrados no *Ensembl Release 90*.
2. Dentre os 41 genes descritos para o cromossomo Y de *C. jacchus*, o número de 22 possui dados referentes a ortólogos, estando quatro distribuídos entre cromossomos sexuais (X e Y) e 18 localizados em diferentes autossomos das espécies *Homo sapiens*, *Pongo abelii*, *Macaca mulata* e *Chlorocebus sabaesus*, evidenciando macrossintenia entre essas espécies.
3. A respeito dos cromossomos sexuais Y das espécies de primatas analisadas, em termos de extensão e quantidade de genes, *C. jacchus* apresenta o menor cromossomo Y e menor quantidade de genes descritos, diferenciando-o dos Catarrhini.
4. A construção dos mapas genéticos utilizando os genes ortólogos revelou a presença de macrossintenia e quebra de colinearidade entre as espécies estudadas, bem como enfatizou uma maior distribuição dos genes presentes no cromossomo Y de *C. jacchus* no cromossomo 7 de *Sus scrofa* e *Macaca mulatta*, cromossomo 26 de *Chlorocebus sabaesus* e cromossomo 15 de *Homo sapiens* e *Pongo abelii*.
5. Foram sugeridas evidências de possíveis translocações e inversões entre um autossomo ancestral e o cromossomo Y em *C. jacchus*, baseadas na ordem dos genes ortólogos visualizada nas espécies de primatas avaliadas.

REFERÊNCIAS BIBLIOGRÁFICAS¹

- BACHTROG, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. **Nature Reviews Genetics**, v. 14, n. 2, p. 113, 2013.
- BAILEY, T. L. et al. MEME SUITE: tools for motif discovery and searching. **Nucleic acids research**, v. 37, n. suppl_2, p. W202-W208, 2009.
- BALCH, W. E. Small GTP-binding proteins in vesicular transport. **Trends in biochemical sciences**, v. 15, n. 12, p. 473-477, 1990.
- BAYLEY, H. Sequencing single molecules of DNA. **Current opinion in chemical biology**, v. 10, n. 6, p. 628-637, 2006.
- BENIRSCHKE, K.; ANDERSON, J. M.; BROWNHILL, L. E. Marrow chimerism in marmosets. **Science**, v. 138, n. 3539, p. 513-515, 1962.
- BIGONI, F. et al. Fluorescence in situ hybridization establishes homology between human and silvered leaf monkey chromosomes, reveals reciprocal translocations between chromosomes homologous to human Y/5, 1/9, and 6/16, and delineates an X1X2Y1Y2/X1X1X2X2 sex-chromosome system. **American journal of physical anthropology**, v. 102, n. 3, p. 315-327, 1997.
- BOHLE, Harry M.; GABALDÓN, Toni. Selection of marker genes using whole-genome DNA polymorphism analysis. **Evolutionary Bioinformatics**, v. 8, p. 161, 2012.
- BOURQUE, G.; PEVZNER, P. A.; TESLER, G. Reconstructing the genomic architecture of ancestral mammals: lessons from human, mouse, and rat genomes. **Genome research**, v. 14, n. 4, p. 507-516, 2004.
- BOURQUE, G. et al. Comparative architectures of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages. **Genome research**, v. 15, n. 1, p. 98-110, 2005.
- BUSO, G. S. C. Marcadores moleculares e análise filogenética. **Embrapa Recursos Genéticos e Biotecnologia. Documentos**, 2005.
- CHARCHAR, F. J. et al. Complex events in the evolution of the human pseudoautosomal region 2 (PAR2). **Genome research**, v. 13, n. 2, p. 281-286, 2003.
- CAPOZZI, O. et al. The 14/15 association as a paradigmatic example of tracing karyotype evolution in New World monkeys. **Chromosoma**, v. 125, n. 4, p. 747-756, 2016.
- CONTRERAS, L. C.; TORRES-MURA, J. C.; SPOTORNO, A. E. The largest known chromosome number for a mammal, in a South American desert rodent. *Experientia*, v. 46, n. 5, p. 506-508, 1990.

¹ De acordo com Associação Brasileira de Normas Técnicas. NBR 6023.

DE ALMEIDA VIEIRA, F. et al. Diversidade genética de *Copernicia prunifera* com o uso de marcadores moleculares ISSR. **Revista Brasileira de Ciências Agrárias**, v. 10, n. 4, 2015.

DE OLIVEIRA, E. H. C.; NEUSSER, M.; MÜLLER, S. Chromosome evolution in new world monkeys (Platyrrhini). **Cytogenetic and genome research**, v. 137, n. 2-4, p. 259-272, 2012.

DEMOULIN, Jean-Baptiste; RENAULD, Jean-Christophe. Interleukin 9 and its receptor: an overview of structure and function. **International reviews of immunology**, v. 16, n. 3-4, p. 345-364, 1998.

DIXIT, R. et al. PHYSICAL MAPPING OF GENOME AND GENES. *J. Biol. Engg. Res. & Rev.*, Vol. 1, n. 1, p. 06-11, 2014.

DIXSON, A. Primate sexuality. Victoria University of Wellington, New Zealand, 2015.

DUMAS, F.; MAZZOLENI, S. Neotropical primate evolution and phylogenetic reconstruction using chromosomal data. **The European Zoological Journal**, v. 84, n. 1, p. 1-18, 2017.

ELLIOTT, D. J. et al. Expression of RBM in the nuclei of human germ cells is dependent on a critical region of the Y chromosome long arm. **Proceedings of the National Academy of Sciences**, v. 94, n. 8, p. 3848-3853, 1997.

FARRÉ, M.; MICHELETTI, D.; RUIZ-HERRERA, A. Recombination rates and genomic shuffling in human and chimpanzee—a new twist in the chromosomal speciation theory. **Molecular biology and evolution**, v. 30, n. 4, p. 853-864, 2012.

FERRETTI, V.; NADEAU, J. H.; SANKOFF, D. Original synteny. In: **Annual Symposium on Combinatorial Pattern Matching**. Springer, Berlin, Heidelberg, 1996. p. 159-167.

FISCHER, M.; THATTE, B. Revisiting an equivalence between maximum parsimony and maximum likelihood methods in phylogenetics. **Bulletin of mathematical biology**, v. 72, n. 1, p. 208-220, 2010.

FORMAN, M. R. et al. Bioinformatics: tools to accelerate population science and disease control research. **American journal of preventive medicine**, v. 38, n. 6, p. 646-651, 2010.

GAO, Sen et al. Determining Sequence Identities: BLAST, Phylogenetic Analysis, and Syntenic Analyses. **Bioinformatics in Aquaculture: Principles and Methods**, 2017.

GODINOT, Marc. Paleocene and Eocene Primates. **The International Encyclopedia of Primatology**, 2017.

GRADA, A.; WEINBRECHT, K. Next-generation sequencing: methodology and application. **The Journal of investigative dermatology**, v. 133, n. 8, p. e11, 2013.

GUT, I. G. New sequencing technologies. **Clinical and Translational Oncology**, v. 15, n. 11, p. 879-881, 2013.

HALL, T.; BIOSCIENCES, I.; CARLSBAD, C. BioEdit: an important software for molecular biology. **GERF Bull Biosci**, v. 2, n. 1, p. 60-61, 2011.

- HARRIS, R. A. et al. Evolutionary genetics and implications of small size and twinning in callitrichine primates. **Proceedings of the National Academy of Sciences**, v. 111, n. 4, p. 1467-1472, 2014.
- HUGHES, J. F.; ROZEN, S. Genomics and genetics of human and primate Y chromosomes. **Annual review of genomics and human genetics**, v. 13, p. 83-108, 2012.
- HUGHES, J. F.; PAGE, D. C. The biology and evolution of mammalian Y chromosomes. **Annual review of genetics**, v. 49, p. 507-527, 2015.
- ISEA, R. The present-day meaning of the word bioinformatics. **Glob J Adv Res**, v. 2, p. 70-73, 2015.
- JIANGTAO, C. et al. MapGene2Chrom, a tool to draw gene physical map based on Perl and SVG languages. **Yi chuan= Hereditas**, v. 37, n. 1, p. 91-97, 2015.
- KEMKEMER, C. et al. Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution. **BMC evolutionary biology**, v. 9, n. 1, p. 84, 2009.
- KIDO, T.; LAU, Yun-Fai C. The rat Tspy is preferentially expressed in elongated spermatids and interacts with the core histones. **Biochemical and biophysical research communications**, v. 350, n. 1, p. 56-67, 2006.
- KIM, H. et al. Comparative physical mapping between *Oryza sativa* (AA genome type) and *O. punctata* (BB genome type). **Genetics**, v. 176, n. 1, p. 379-390, 2007.
- KONKEL, M. K. et al. Discovery of a new repeat family in the *Callithrix jacchus* genome. **Genome research**, v. 26, n. 5, p. 649-659, 2016.
- KUROIWA, A. et al. The process of a Y-loss event in an XO/XO mammal, the Ryukyu spiny rat. **Chromosoma**, v. 119, n. 5, p. 519-526, 2010.
- LAU, Y. C. Gonadoblastoma, testicular and prostate cancers, and the TSPY gene. **The American Journal of Human Genetics**, v. 64, n. 4, p. 921-927, 1999.
- LI, G. et al. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. **Genome research**, v. 23, n. 9, p. 1486-1495, 2013.
- LU, D. Y. et al. Cancer bioinformatics, its impacts on cancer therapy. **Metabolomics**, v. 5, n. 2, p. e133, 2015.
- MALUKIEWICZ, J. et al. Phylogeny of the *jacchus* group of *Callithrix* marmosets based on complete mitochondrial genomes. **American journal of physical anthropology**, v. 162, n. 1, p. 157-169, 2017.
- MANSFIELD, K. Marmoset models commonly used in biomedical research. **Comparative medicine**, v. 53, n. 4, p. 383-392, 2003.
- MARSH, A. Evolution of the Y-Chromosome in Primates. 2014.
- MARTIN, R. D. Primates. **Current Biology**, v. 22, n. 18, 2012.
- MCGINN, S.; GUT, I. G. DNA sequencing—spanning the generations. **New biotechnology**, v. 30, n. 4, p. 366-372, 2013.

- METZKER, M. L. Sequencing technologies—the next generation. **Nature reviews genetics**, v. 11, n. 1, p. 31, 2010.
- MITCHELL, J. F.; LEOPOLD, D. A. The marmoset monkey as a model for visual neuroscience. **Neuroscience research**, v. 93, p. 20-46, 2015.
- MOORE, J. H.; ASSELBERGS, F. W.; WILLIAMS, S. M. Bioinformatics challenges for genome-wide association studies. **Bioinformatics**, v. 26, n. 4, p. 445-455, 2010.
- MÜLLER, S. Primate chromosome evolution. **Genomic Disorders**, p. 133-152, 2006.
- MURPHY, W. J.; STANYON, R.; O'BRIEN, S. J. Evolution of mammalian genome organization inferred from comparative gene mapping. **Genome biology**, v. 2, n. 6, p. reviews0005. 1, 2001.
- MURPHY, W. J. et al. Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. **Science**, v. 309, n. 5734, p. 613-617, 2005.
- ORSI, A. et al. Overview of the marmoset as a model in nonclinical development of pharmaceutical products. **Regulatory Toxicology and Pharmacology**, v. 59, n. 1, p. 19-27, 2011.
- PARKS, S. L.; GOLDMAN, N. Maximum likelihood inference of small trees in the presence of long branches. **Systematic biology**, v. 63, n. 5, p. 798-811, 2014.
- PERELMAN, P. et al. A molecular phylogeny of living primates. **PLoS genetics**, v. 7, n. 3, p. e1001342, 2011.
- RAPPOLD, G. A. The pseudoautosomal regions of the human sex chromosomes. **Human genetics**, v. 92, n. 4, p. 315-324, 1993.
- RAUT, S. A.; SATHE, S. R.; RAUT, Adarsh. Bioinformatics: Trends in gene expression analysis. In: **Bioinformatics and Biomedical Technology (ICBBT), 2010 International Conference on**. IEEE, 2010. p. 97-100.
- RAZA, K. Application of data mining in bioinformatics. **arXiv preprint arXiv:1205.1125**, 2012.
- RYLANDS, ANTHONY B.; COIMBRA-FILHO, A. F.; MITTERMEIER, R. A. Systematics, geographic distribution, and some notes on the conservation status of the Callitrichidae. *Marmosets and tamarins: Systematics, behaviour, and ecology*, p. 11-77, 1993.
- SANGER, F.; NICKLEN, S.; COULSON, A. R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the national academy of sciences**, v. 74, n. 12, p. 5463-5467, 1977.
- SCHARASCHKIN, T. Why is there no genetics in phylogenetics: misconceptions, clarification and solutions. In: **Proceedings of the Royal Society of Queensland**. Royal Society of Queensland, 2013. p. 63-67.
- SCHIBLER, L. et al. Comparative gene mapping: a fine-scale survey of chromosome rearrangements between ruminants and humans. **Genome Research**, v. 8, n. 9, p. 901-915, 1998.
- SHENDURE, J.; AIDEN, E. L. The expanding scope of DNA sequencing. **Nature biotechnology**, v. 30, n. 11, p. 1084, 2012.

- SHERLOCK, J. K. et al. Homologies between human and marmoset (*Callithrix jacchus*) chromosomes revealed by comparative chromosome painting. **Genomics**, v. 33, n. 2, p. 214-219, 1996.
- SIEVERS, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular systems biology**, v. 7, n. 1, p. 539, 2011.
- TAMURA, K. et al. MEGA6: molecular evolutionary genetics analysis version 6.0. **Mol. Biol. Evol.**, v. 30, p. 2725–2729, 2013.
- TAN, Eng-King. LINGO1 and essential tremor: linking the shakes. Linking LINGO1 to essential tremor. 2010.
- TANG, Haibao et al. Synteny and collinearity in plant genomes. **Science**, v. 320, n. 5875, p. 486-488, 2008.
- STANYON, R. et al. Primate chromosome evolution: ancestral karyotypes, marker order and neocentromeres. **Chromosome Research**, v. 16, n. 1, p. 17-39, 2008.
- The Marmoset Genome Sequencing And Analysis Consortium, The Marmoset Genome et al. The common marmoset genome provides insight into primate biology and evolution. **Nature genetics**, v. 46, n. 8, p. 850, 2014.
- THOMPSON, J. D. et al. Multiple sequence alignment using ClustalW and ClustalX. **Current protocols in bioinformatics**, p. 2.3. 1-2.3. 22, 2002.
- TORRES, L. B. et al. The use of new world primates for biomedical research: an overview of the last four decades. **American journal of primatology**, v. 72, n. 12, p. 1055-1061, 2010.
- STANYON, R. et al. Evolutionary molecular cytogenetics of catarrhine primates: past, present and future. **Cytogenetic and genome research**, v. 137, n. 2-4, p. 273-284, 2012.
- VAN DIJK, E. L. et al. Ten years of next-generation sequencing technology. **Trends in genetics**, v. 30, n. 9, p. 418-426, 2014.
- VAWTER, M. P.; HARVEY, P. D.; DELISI, L. E. Dysregulation of X-linked gene expression in Klinefelter's syndrome and association with verbal cognition. **American Journal of Medical Genetics Part B: Neuropsychiatric Genetics**, v. 144, n. 6, p. 728-734, 2007.
- VERMEESCH, J. R. et al. The IL-9 receptor gene, located in the Xq/Yq pseudoautosomal region, has an autosomal origin, escapes X inactivation and is expressed from the Y. **Human molecular genetics**, v. 6, n. 1, p. 1-8, 1997.
- VILARIÑO-GÜELL, C. et al. LINGO1 and LINGO2 variants are associated with essential tremor and Parkinson disease. **neurogenetics**, v. 11, n. 4, p. 401-408, 2010.
- XU, Yiqing et al. VGSC: a web-based Vector Graph toolkit of genome Synteny and Collinearity. **BioMed research international**, v. 2016, 2016.
- WATERS, P. D.; WALLIS, M. C.; GRAVES, J. A. Marshall. Mammalian sex—origin and evolution of the Y chromosome and SRY. In: **Seminars in cell & developmental biology**. Academic Press, 2007. p. 389-400.

WEI, Y. et al. Comparative physical genome mapping of malaria vectors *Anopheles sinensis* and *Anopheles gambiae*. **Malaria journal**, v. 16, n. 1, p. 235, 2017.

YANG, Z.; RANNALA, B. Molecular phylogenetics: principles and practice. **Nature Reviews Genetics**, v. 13, n. 5, p. 303, 2012.

APÊNDICE A- QUADROS DOS GENES

Quadro 1 - Genes anotados para construção do mapa genético do cromossomo Y de *Callithrix jacchus*, bem como para investigação das relações sintênicas, enfatizando o ID disponível no NCBI, o nome do gene encontrado no NCBI (<http://www.ncbi.nlm.nih.gov>) e no *Ensembl Release 90* (<http://www.ensembl.org>), bem como o código, tipo, tamanho em bases e localização dessa sequência gênica disponível no *Ensembl Release 90*.

Nº	ID Gene (NCBI)	Nome do gene	Código (<i>Ensembl Release 90</i>)	Tipo do gene (<i>Ensembl Release 90</i>)	Tamanho (b) (<i>Ensembl Release 90</i>)	Localização (<i>Ensembl Release 90</i>)
1	100405601	LINGO1	ENSCJAG00000011159	protein_coding	85669	Chromosome Y: 822747-908415 forward strand
2	100397261	GTPBP6	ENSCJAG00000013056	protein_coding	7402	Chromosome Y: 14554-21955 reverse strand
3	-	-	ENSCJAG00000025300	miRNA	75	Chromosome Y: 616341-616415 reverse strand
4	100400516	RBMV	ENSCJAG00000033808	protein_coding	16838	Chromosome Y: 520700-537537 reverse strand
5	-	-	ENSCJAG00000023103 (sem ortólogo)	Pseudo	1026	Chromosome Y: 170054-171079 reverse strand
6	-	-	ENSCJAG00000011164	protein_coding	5952	Chromosome Y: 611783-617734 forward strand
7	-	-	ENSCJAG00000034662 (sem ortólogo)	Pseudo	777	Chromosome Y: 246191-246967 reverse strand
8	-	-	ENSCJAG00000015432 (sem ortólogo)	Pseudo	642	Chromosome Y: 195441-196082 reverse strand
9	-	-	ENSCJAG00000034791	protein_coding	2839	Chromosome Y: 410879-413717 forward strand
10	-	-	ENSCJAG00000032219 (ortólogo de sp inviável)	protein_coding	792	Chromosome Y: 342818-343609 forward strand
11	-	-	ENSCJAG00000011168	protein_coding	1998	Chromosome Y: 600696-602693 forward strand
12	-	-	ENSCJAG00000034293	protein_coding	1207	Chromosome Y: 417973-419179 reverse strand
13	-	-	ENSCJAG00000033485	protein_coding	1171	Chromosome Y: 578422-579592 forward strand
14	100402952	SCAPER	ENSCJAG00000010868	protein_coding	531668	Chromosome Y: 1691591-2223258 forward strand
15	100405972	PEAK1	ENSCJAG00000011124	protein_coding	60475	Chromosome Y: 1388111-1448585 forward strand
16	100406340	PSTPIP1	ENSCJAG00000011058	protein_coding	47613	Chromosome Y: 1523596-1571208 reverse strand
17	100401619	HMG20A	ENSCJAG00000011131	protein_coding	82873	Chromosome Y: 1053082-1135954 reverse strand
18	-	-	ENSCJAG00000032048 (sem ortólogo)	pseudogene	1132	Chromosome Y: 1839049-1840180 forward strand
19	100401992	TSPAN3	ENSCJAG00000011087	protein_coding	24883	Chromosome Y: 1487192-1512074 forward strand
20	100406707	RCN2	ENSCJAG00000010987	protein_coding	18616	Chromosome Y: 1616816-1635431 reverse strand
21	-	-	ENSCJAG00000011052 (sem ortólogo)	pseudogene	797	Chromosome Y: 1586217-1587013 forward strand
22	-	Metazoa_SRP	ENSCJAG00000037657	misc_RNA	293	Chromosome Y: 1588665-1588957 forward strand
23	-	U6	ENSCJAG00000026644 (poucos ortólogos)	snRNA	104	Chromosome Y: 1642703-

						1642806 forward strand
24	100404521	ETFA	ENSCJAG00000010791	protein_coding	93074	Chromosome Y: 2269998-2363071 forward strand
25	100404151	ISL2	ENSCJAG00000010855	protein_coding	8088	Chromosome Y: 2228289-2236376 reverse strand
26	100404877	TMEM266	ENSCJAG00000010771	protein_coding	150163	Chromosome Y: 2374619-2524781 reverse strand
27	100408034	IL9R	ENSCJAG00000008510	protein_coding	12847	Chromosome Y: 2819940-2832786 forward strand
28	100407430	VAMP7	ENSCJAG00000016289	protein_coding	60270	Chromosome Y: 2745956-2806225 forward strand

Quadro 2 - Genes anotados para construção do mapa genético de *Homo sapiens* e investigação das relações sintênicas, enfatizando o nome do gene, bem como o código, o tipo, o tamanho em bases (b) e a localização do gene, conforme disponível no *Ensembl Release 90*.

N°	Nome	Código (Ensembl Release 90)	Tipo do gene	Tamanho	Localização	
1	LINGO1	ENSG00000169783	protein_coding	207874	Chromosome 15: 77613027-77820900 reverse strand	
2	GTPBP6	ENSG00000178605	protein_coding	14291	Chromosome X: 304529-318819 reverse strand	
3	-	-	-	-	-	
4	A	RBM1Y1F	ENSG00000169800	Protein coding	14441	Chromosome Y: 22168542-22182982 reverse strand.
	B	RBM1Y1J	ENSG00000226941	protein_coding	14421	Chromosome Y: 22403461-22417881 forward strand
	C	RBM1Y1A1	ENSG00000234414	protein_coding	24805	Chromosome Y: 21511372-21549326 forward strand
	D	RBM1Y1E	ENSG00000242389	protein_coding	14450	Chromosome Y: 21903618-21918067 reverse strand
	E	RBM1Y1B	ENSG00000242875	protein_coding	37757	Chromosome Y: 21511338-21549094 forward strand
	F	RBM1Y1D	ENSG00000244395	protein_coding	14451	Chromosome Y: 21880076-21894526 reverse strand
5	-	-	-	-	-	
6	NEIL1	ENSG00000140398	protein_coding	10160	Chromosome 15: 75346955-75357114 forward strand	
7	-	-	-	-	-	
8	-	-	-	-	-	
9	A	TSPY2	ENSG00000168757	protein_coding	2797	Chromosome Y: 6246223-6249019 forward strand
	B	TSPY3	ENSG00000228927	protein_coding	2803	Chromosome Y: 9398421-9401223 forward strand
	C	TSPY8	ENSG00000229549	protein_coding	2803	Chromosome Y: 9357797-9360599 forward strand
	D	TSPY4	ENSG00000233803	protein_coding	2821	Chromosome Y: 9,337,464-9,340,284 forward strand
	E	TSPY10	ENSG00000236424	protein_coding	2803	Chromosome Y: 9527880-9530682 forward strand
	F	TSPY1	ENSG00000258992	protein_coding	2794	Chromosome Y: 9466955-9469748 forward strand
10	-	-	-	-	-	
11	COMMD4	ENSG00000140365	protein_coding	7334	Chromosome 15: 75335891-75343224 forward strand	

12	-	-	-	-	-
13	-	-	-	-	-
14	SCAPER	ENSG00000140386	protein_coding	557541	Chromosome 15: 76347904-76905444 reverse strand
15	PEAK1	ENSG00000173517	protein_coding	319489	Chromosome 15: 77100656-77420144 reverse strand
16	PSTPIP1	ENSG00000140368	protein_coding	43974	Chromosome 15: 76993359-77037332 forward strand
17	HMG20A	ENSG00000140382	protein_coding	65196	Chromosome 15: 77420412-77485607 forward strand
18	-	-	-	-	-
19	TSPAN3	ENSG00000140391	protein_coding	42581	Chromosome 15: 77041404-77083984 reverse strand
20	RCN2	ENSG00000117906	protein_coding	22774	Chromosome 15: 76931619-76954392 forward strand
21	-	-	-	-	-
22	Metazoa_SRP	ENSG00000243365	Misc RNA	295	Chromosome 15: 76976915-76977210 reverse strand
23	RNU6-1029P	ENSG00000206836	snRNA	107	Chromosome 2: 197447745-197447851 forward strand
24	ETFA	ENSG00000140374	protein_coding	96118	Chromosome 15: 76215355-76311472 reverse strand
25	ISL2	ENSG00000159556	protein_coding	5753	Chromosome 15: 76336724-76342476 forward strand
26	TMEM266	ENSG00000169758	protein_coding	169285	Chromosome 15: 76059837-76229121 forward strand
27	IL9R	ENSG00000124334	protein_coding	13237	Chromosome X: 155997581-156010817 forward strand
28	VAMP7	ENSG00000124333	protein_coding	62477	Chromosome X: 155,881,293-155,943,769 forward strand

Quadro 3 - Genes anotados para construção do mapa genético de *Pongo abelii* e investigação das relações sintênicas, enfatizando o nome do gene, bem como o código, o tipo, o tamanho em bases (b) e a localização do gene, conforme disponível no *Ensembl Release 90*.

Nº	Nome	Nome ensembl	Tipo do gene	Tamanho	Localização
1	-	-	-	-	-
2	GTPBP6	ENSPPYG00000020075	protein_coding	9241	Chromosome X: 41503-50743 reverse strand
3	-	ENSPPYG00000021136	miRNA	74	Chromosome 15: 72821876-72821950 reverse strand
4	-	-	-	-	-
5	-	-	-	-	-
6	NEIL1	ENSPPYG00000006658	protein_coding	7792	Chromosome 15: 72815514-72823305 forward strand
7	-	-	-	-	-
8	-	-	-	-	-
9	-	-	-	-	-
10	-	-	-	-	-
11	COMMD4	ENSPPYG00000006657	protein_coding	6145	Chromosome 15: 72803724-72809868 forward strand
12	-	-	-	-	-
13	-	ENSPPYG00000006668	Protein coding	1440	Chromosome 15: 73284164-73285604 forward strand
14	SCAPER	ENSPPYG00000006677	protein_coding	566121	Chromosome 15: 73866368-74432488 reverse strand
15	PEAK1	ENSPPYG00000006681	protein_coding	67148	Chromosome 15: 74649890-74717037 reverse strand
16	PSTPIP1	ENSPPYG00000006679	protein_coding	41844	Chromosome 15: 74529718-74571561 forward strand

17	HMG20A	ENSPPYG00000006683	protein_coding	67981	Chromosome 15: 74977679-75045659 forward strand
18	-	-	-	-	-
19	TSPAN3	ENSPPYG00000006680	protein_coding	26967	Chromosome 15: 74580341-74607307 reverse strand
20	RCN2	ENSPPYG00000006678	protein_coding	18267	Chromosome 15: 74466204-74484470 forward strand
21	-	-	-	-	-
22	-	-	-	-	-
23	RNU6-1029P	ENSPPYG00000025781	snRNA	106	Chromosome 2b: 88423432-88423538 forward strand
24	ETFA	ENSPPYG00000006674	protein_coding	102952	Chromosome 15: 73727122-73830073 reverse strand
25	ISL2	ENSPPYG00000006676	protein_coding	5724	Chromosome 15: 73854876-73860599 forward strand
26	TMEM266	ENSPPYG00000006673	protein_coding	147064	Chromosome 15: 73572115-73719178 forward strand
27	IL9R	ENSPPYG00000020907	protein_coding	11656	Chromosome X: 156169561-156181216 forward strand
28	VAMP7	ENSPPYG00000020906	protein_coding	57977	Chromosome X: 156051603-156109579 forward strand

Quadro 4 - Genes anotados para construção do mapa genético de *Macaca mulatta* e investigação das relações sintênicas, enfatizando o nome do gene, bem como o código, o tipo, o tamanho em bases (b) e a localização do gene, conforme disponível no *Ensembl Release 90*.

N°	Nome	Nome ensembl	Tipo do gene	Tamanho	Localização
1	LINGO1	ENSMMUG00000001782	protein_coding	204659	Chromosome 7: 55091442-55296100 reverse strand
2	-	-	-	-	-
3	mml-mir-631	ENSMMUG00000027217	miRNA	75	Chromosome 7: 52678504-52678578 reverse strand
4	-	ENSMMUG00000039261	Protein coding	8310	Chromosome Y: 10104257-10112567 forward strand
5	-	-	-	-	-
6	NEIL1	ENSMMUG00000013079	protein_coding	8337	Chromosome 7: 52671784-52680120 forward strand
7	-	-	-	-	-
8	-	-	-	-	-
9	A	ENSMMUG00000038984	protein_coding	1813	Chromosome Y: 9880133-9881945 reverse strand
	B	ENSMMUG00000041055	Protein coding	2800	Chromosome Y: 6306755-6309555 reverse strand
	C	ENSMMUG00000041274	protein_coding	4078	Chromosome Y: 9927166-9931243 reverse strand
	D	ENSMMUG00000043551	protein_coding	1937	Chromosome Y: 9912910-9914846 reverse strand
	E	ENSMMUG00000048285	Protein coding	1937	Chromosome Y: 9896509-9898445 reverse strand
	F	ENSMMUG00000049217	protein_coding	1972	Chromosome Y: 10067088-10069059 forward strand
10	-	-	-	-	-
11	COMMD4	ENSMMUG00000047011	protein_coding	7683	Chromosome 7: 52663323-52671005 forward strand
12	-	-	-	-	-
13	-	ENSMMUG00000043684	Protein coding	2403	Chromosome 7: 56134197-56136600 forward strand
14	SCAPER	ENSMMUG00000011047	protein_coding	556280	Chromosome 7: 53797192-54353471 reverse strand
15	PEAK1	ENSMMUG00000042378	protein_coding	142994	Chromosome 7: 54585968-54728961 reverse strand
16	PSTPIP1	ENSMMUG00000010171	protein_coding	42460	Chromosome 7: 54471788-54514247 forward strand

17	HMG20A	ENSMMUG00000007430	protein_coding	67788	Chromosome 7: 54891331-54959118 forward strand
18	-	-	-	-	-
19	TSPAN3	ENSMMUG00000010172	protein_coding	25392	Chromosome 7: 54524008-54549399 reverse strand
20	RCN2	ENSMMUG00000008213	protein_coding	20299	Chromosome 7: 54407751-54428049 forward strand
21	-	-	-	-	-
22	Metazoa_SRP	ENSMMUG00000049200	misc RNA	286	Chromosome 7: 54453714-54454000 reverse strand
23	-	-	-	-	-
24	ETFA	ENSMMUG00000011045	protein_coding	102198	Chromosome 7: 53658723-53760920 reverse strand
25	-	-	-	-	-
26	-	-	-	-	-
27	-	ENSMMUG00000002016	Protein coding	10081	Chromosome X: 149126657-149136738 forward strand
28	VAMP7	ENSMMUG00000002015	protein_coding	59122	Chromosome X: 149035967-149095088 forward strand

Quadro 5 - Genes anotados para construção do mapa genético de *Chlorocephus sabaeus* e investigação das relações sintênicas, enfatizando o nome do gene, bem como o código, o tipo, o tamanho em bases (b) e a localização do gene, conforme disponível no *Ensembl Release 90*.

Nº	Nome	Nome ensembl	Tipo do gene	Tamanho	Localização
1	LINGO1	ENSCSAG00000019731	protein_coding	1863	Chromosome 26: 5688293-5690155 forward strand
2	-	-	-	-	-
3	-	ENSCSAG00000023682	miRNA	74	Chromosome 26: 8076245-8076319 forward strand
4	-	ENSCSAG00000018310	Protein coding	10860	Chromosome Y: 4036571-4047431 forward strand
5	-	-	-	-	-
6	NEIL1	ENSCSAG00000014386	protein_coding	7814	Chromosome 26: 8074310-8082123 reverse strand
7	-	-	-	-	-
8	-	-	-	-	-
9	-	-	-	-	-
10	-	-	-	-	-
11	COMMD4	ENSCSAG00000014380	protein_coding	2430	Chromosome 26: 8086872-8089301 reverse strand
12	-	ENSCSAG00000018306	Protein coding	18077	Chromosome Y: 5350554-5368631 reverse strand
13	-	ENSCSAG00000011652	protein_coding	4974	Chromosome 9: 29966683-29971656 reverse strand
14	SCAPER	ENSCSAG00000014700	Protein coding	401448	Chromosome 26: 6533124-6934572 forward strand
15	PEAK1	ENSCSAG00000014728	protein_coding	59646	Chromosome 26: 6125081-6184726 forward strand
16	PSTPIP1	ENSCSAG00000014714	protein_coding	42325	Chromosome 26: 6260349-6302673 reverse strand
17	HMG20A	ENSCSAG00000014734	protein_coding	27584	Chromosome 26: 5821105-5848688 reverse strand
18	-	-	-	-	-
19	TSPAN3	ENSCSAG00000014721	protein_coding	26287	Chromosome 26: 6225610-6251896 forward strand
20	RCN2	ENSCSAG00000014707	Protein coding	18583	Chromosome 26: 6347539-6366122 reverse strand
21	-	-	-	-	-

22	Metazoa_S RP	ENSCSAG00000020669	misc RNA	292	Chromosome 26: 6320479-6320771 forward strand
23	-	-	-	-	-
24	ETFA	ENSCSAG00000014683	protein_coding	105466	Chromosome 26: 6970584-7076049 forward strand
25	ISL2	ENSCSAG00000014692	protein_coding	4937	Chromosome 26: 6940850-6945786 reverse strand
26	TMEM266	ENSCSAG00000014673	protein_coding	68632	Chromosome 26: 7086016-7154647 reverse strand
27	IL9R	ENSCSAG00000006485	protein_coding	13996	Chromosome X: 130006899-130020894 forward strand
28	VAMP7	ENSCSAG00000006492	protein_coding	47074	Chromosome X: 129937046-129984119 forward strand

Quadro 6 - Genes anotados para construção do mapa genético de *Sus scrofa* e investigação das relações sintênicas, enfatizando o nome do gene, bem como o código, o tipo, o tamanho em bases (b) e a localização do gene, conforme disponível no *Ensembl Release 90*.

Nº	Nome	Nome ensembl	Tipo do gene	Tamanho	Localização
1	LINGO1	ENSSSCG00000001872	protein_coding	205538	Chromosome 7: 57684589-57890126 reverse strand
2	-	ENSSSCG000000033858	protein_coding	11991	Chromosome Y: 6984332-6996322 reverse strand
3	-	-	-	-	-
4	-	ENSSSCG000000034141	protein_coding	9257	Chromosome Y: 41171996-41181252 reverse strand
5	-	-	-	-	-
6	NEIL1	ENSSSCG00000001882	protein_coding	9251	Chromosome 7: 58259756-58269006 reverse strand
7	-	-	-	-	-
8	-	-	-	-	-
9	-	ENSSSCG000000023819	protein_coding	1981	Chromosome Y: 40190400-40192380 reverse strand
10	-	-	-	-	-
11	COMMD4	ENSSSCG00000001884	protein_coding	6483	Chromosome 7: 58276959-58283441 reverse strand
12	-	-	-	-	-
13	-	ENSSSCG000000005643	protein_coding	19666	Chromosome 1: 268701511-268721176 reverse strand
14	SCAPER	ENSSSCG00000001865	protein_coding	430234	Chromosome 7: 56582136-57012369 reverse strand
15	PEAK1	ENSSSCG00000001869	protein_coding	279831	Chromosome 7: 57223818-57503648 reverse strand
16	PSTPIP1	ENSSSCG00000001867	protein_coding	43995	Chromosome 7: 57102757-57146751 forward strand
17	HMG20A	ENSSSCG00000001871	protein_coding	65902	Chromosome 7: 57504259-57570160 forward strand
18	-	-	-	-	-
19	TSPAN3	ENSSSCG00000001868	protein_coding	28600	Chromosome 7: 57158916-57187515 reverse strand
20	RCN2	ENSSSCG00000001866	protein_coding	17715	Chromosome 7: 57041033-57058747 forward strand
21	-	-	-	-	-
22	-	-	-	-	-
23	-	ENSSSCG000000040320	snRNA	106	Chromosome 13: 120922737-120922843 forward strand
24	ETFA	ENSSSCG00000001862	protein_coding	100513	Chromosome 7: 56440620-56541132 reverse strand
25	ISL2	ENSSSCG000000033348	protein_coding	5841	Chromosome 7: 56571419-56577259 forward strand
26	TMEM266	ENSSSCG00000001863	protein_coding	129158	Chromosome 7: 56311522-56440679 forward strand
27	IL9R	ENSSSCG00000007973	protein_coding	8417	Chromosome 3: 41582065-41590481 forward strand

28	VAMP7	ENSSSCG00000029615	protein_coding	53959	Chromosome X: 125863273-125917231 forward strand
----	-------	--------------------	----------------	-------	--

Quadro 7 - Genes ortólogos do cromossomo Y de *Callithrix jacchus* anotados a partir do Ensemble Release 90 descritos em *Homo sapiens*, *Pongo abelii*, *Macaca mulata*, *Chlorocebus sabaeus* e *Sus scrofa*.

Nº	Nome do gene (<i>Callithrix jacchus</i>)	<i>Homo sapiens</i>	<i>Pongo abelii</i>	<i>Macaca mulatta</i>	<i>Chlorocebus sabaeus</i>	<i>Sus scrofa</i>	Nome do gene
1	ENSCJAG000000111 59	ENSG00000169783 (Chromosome 15)	-	ENSMMLUG0000000178 2 (Chromosome 7)	ENSCSAG000000019 731 (Chromosome 26)	ENSSSCG000000001 872 (Chromosome 7)	LINGO1
2	ENSCJAG000000130 56	ENSG00000178605 (Chromosome X)	ENSPPYG0000002007 5 (Chromosome X); #	-	-	ENSSSCG0000000338 58 (Chromosome Y) #	GTPBP6#
3	ENSCJAG000000253 00	-	ENSPPYG0000002113 6 (Chromosome 15)	ENSMMLUG0000002721 7 (Chromosome 7)	ENSCSAG000000023 682 (Chromosome 26)	-	
4	ENSCJAG000000338 08	ENSG00000169800 (Chromosome Y); ENSG00000226941 (Chromosome Y); ENSG00000234414 (Chromosome Y); ENSG00000242389 (Chromosome Y); ENSG00000242875 (Chromosome Y);	-	ENSMMLUG0000003926 1 (Chromosome Y) #	ENSCSAG000000018 310 (Chromosome Y) #	ENSSSCG000000034 141 (Chromosome Y) #	RBMY
		ENSG00000244395 (Chromosome Y)					
5	ENSCJAG000000231 03 (sem ortólogo)	-	-	-	-	-	

6	ENSCJAG000000111 64 #	ENSG000000140398 (Chromosome 15)	ENSPPYG0000000665 8 (Chromosome 15)	ENSMMUG0000001307 9 (Chromosome 7)	ENSCSAG00000014 386 (Chromosome 26)	ENSSSCG000000001 882 (Chromosome 7)	NEIL1#
7	ENSCJAG000000346 62 (sem ortólogo)	-	-	-	-	-	
8	ENSCJAG000000154 32 (sem ortólogo)	-	-	-	-	-	
9	ENSCJAG000000347 91	ENSG000000168757 (Chromosome Y); ENSG000000228927 (Chromosome Y); ENSG000000229549 (Chromosome Y); ENSG000000233803 (Chromosome Y); ENSG000000236424 (Chromosome Y);	-	ENSMMUG0000003898 4 (Chromosome Y); ENSMMUG000000410 5 ENSMMUG000000412 7 4 (Chromosome Y); ENSMMUG0000004355 1 (Chromosome Y); ENSMMUG0000004828 5 (Chromosome Y); ENSMMUG0000004921 7 (Chromosome Y)	-	ENSSSCG000000023 819 (Chromosome Y)	
10	ENSCJAG000000322 19 (ortólogo de sp inviável)	ENSG000000258992 (Chromosome Y)	-	-	-	-	

11	ENSCJAG000000111 68	ENSG000000140365 (Chromosome 15)	ENSPPYG0000000665 7 (Chromosome 15)	ENSMMLUG0000004701 1 (Chromosome 7)	ENSCSAG000000014 380 (Chromosome 26)	ENSSSCG000000001 884 (Chromosome 7)	
12	ENSCJAG000000342 93	-	-	-	ENSCSAG00000018 306 (Chromosome Y)	-	
13	ENSCJAG000000334 85	-	ENSPPYG0000000666 8 (Chromosome 15)	ENSMMLUG0000004368 4 (Chromosome 7)	ENSCSAG00000011 652 (Chromosome 9)	ENSSSCG000000005 643 (Chromosome 1)	
14	ENSCJAG000000108 68	ENSG000000140386 (Chromosome 15)	ENSPPYG0000000667 7 (Chromosome 15)	ENSMMLUG0000001104 7 (Chromosome 7)	ENSCSAG00000014 700 (Chromosome 26)	ENSSSCG000000001 865 (Chromosome 7)	SCAPER
15	ENSCJAG000000111 24	ENSG000000173517 (Chromosome 15)	ENSPPYG0000000668 1 (Chromosome 15)	ENSMMLUG0000004237 8 (Chromosome 7)	ENSCSAG00000014 728 (Chromosome 26)	ENSSSCG000000001 869 (Chromosome 7)	PEAK1
16	ENSCJAG000000110 58	ENSG000000140368 (Chromosome 15)	ENSPPYG0000000667 9 (Chromosome 15)	ENSMMLUG0000001017 1 (Chromosome 7)	ENSCSAG00000014 714 (Chromosome 26)	ENSSSCG000000001 867 (Chromosome 7)	PSTPIP1
17	ENSCJAG000000111 31	ENSG000000140382 (Chromosome 15)	ENSPPYG0000000668 3 (Chromosome 15)	ENSMMLUG0000000743 0 (Chromosome 7)	ENSCSAG00000014 734 (Chromosome 26)	ENSSSCG000000001 871 (Chromosome 7)	HMG20A
18	ENSCJAG000000320 48 (sem ortólogo)	-	-	-	-	-	-
19	ENSCJAG000000110 87	ENSG000000140391 (Chromosome 15)	ENSPPYG0000000668 0 (Chromosome 15)	ENSMMLUG0000001017 2 (Chromosome 7)	ENSCSAG00000014 721 (Chromosome 26)	ENSSSCG000000001 868 (Chromosome 7)	TSPAN3
20	ENSCJAG000000109 87	ENSG000000117906 (Chromosome 15)	ENSPPYG0000000667 8 (Chromosome 15)	ENSMMLUG00000000821 3 (Chromosome 7)	ENSCSAG00000014 707 (Chromosome 26)	ENSSSCG000000001 866 (Chromosome 7)	RCN2

21	ENSCJAG000000110 52 (sem ortólogo)	-	-	-	-	-	-	-	-
22	ENSCJAG0000000376 57	ENSG000000243365 (Chromosome 15)	-	ENSMMLUG00000004920 0 (Chromosome 7)	ENSCSAG0000000020 669 (Chromosome 26)	-	-	-	-
23	ENSCJAG0000000266 44 (poucos ortólogos)	ENSG000000206836 (Chromosome 2)	ENSPPYG00000002578 1 (Chromosome 2b)	-	-	ENSSSCG0000000040 320 (Chromosome 13)	-	-	-
24	ENSCJAG0000000107 91	ENSG000000140374 (Chromosome 15)	ENSPPYG00000000667 4 (Chromosome 15)	ENSMMLUG00000001104 5 (Chromosome 7)	ENSCSAG0000000014 683 (Chromosome 26)	ENSSSCG0000000001 862 (Chromosome 7)	ETFA	-	-
25	ENSCJAG0000000108 55	ENSG000000159556 (Chromosome 15)	ENSPPYG00000000667 6 (Chromosome 15)	-	ENSCSAG0000000014 692 (Chromosome 26)	ENSSSCG0000000033 348 (Chromosome 7)	ISL2	-	-
26	ENSCJAG0000000107 71	ENSG000000169758 (Chromosome 15)	ENSPPYG00000000066 7 3 (Chromosome 15)	-	ENSCSAG0000000014 673 (Chromosome 26)	ENSSSCG0000000001 863 (Chromosome 7)	TMEM266	-	-
27	ENSCJAG0000000085 10	ENSG000000124334 (Chromosome X)	ENSPPYG00000002090 7 (Chromosome X)	ENSMMLUG00000000201 6 (Chromosome X)	ENSCSAG0000000006 485 (Chromosome X)	ENSSSCG0000000007 973 (Chromosome 3)	-	-	-
28	ENSCJAG0000000162 89	ENSG000000124333 (Chromosome X)	ENSPPYG00000002090 6 (Chromosome X)	ENSMMLUG00000000201 5 (Chromosome X)	ENSCSAG0000000006 492 (Chromosome X)	ENSSSCG0000000029 615 (Chromosome X)	VAMP7	-	-

Quadro 8 – Genes anotados para construção do mapa genético do cromossomo Y de *C. jacchus* segundo o NCBI.

N°	Nome/Gene ID	Descrição	Localização	Siglas/Aliases	tam
1	LOC108589990 ID: 108589990	uncharacterized LOC108589990 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2658429..2659525)	-	1097 b
2	LOC108589989 ID: 108589989	uncharacterized LOC108589989 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2273388..2275518)	-	2131 b
3	LOC108589988 ID: 108589988	uncharacterized LOC108589988 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1541253..1544137, complement)	-	2885 b
4	LOC108589987 ID: 108589987	interferon-induced transmembrane protein 3 pseudogene [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1693830..1694661)	-	832 b
5	LOC108589986 ID: 108589986	60S ribosomal protein L29 pseudogene [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2078901..2079501)	-	601 b
6	LOC108589985 ID: 108589985	RNA-binding motif protein, Y chromosome, family 1 member B-like [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (549340..553539, complement)	-	4200 b
7	LOC103790417 ID: 103790417	uncharacterized LOC103790417 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1361325..1380101, complement)	-	18777 b
8	LOC103790416 ID: 103790416	uncharacterized LOC103790416 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2527507..2580062)	-	52556 b
9	LOC103790415 ID: 103790415	putative Polycomb group protein ASXL1 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1838121..1840200)	-	2080 b
10	LOC103790413 ID: 103790413	uncharacterized LOC103790413 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1047384..1057668)	-	10285 b
11	LOC103790412 ID: 103790412	uncharacterized LOC103790412 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (715412..792712)	-	77301 b
12	LOC103790409 ID: 103790409	uncharacterized LOC103790409 [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (680134..690153, complement)	-	10,020 b
13	LOC103790408 ID: 103790408	golgin subfamily A member 2-like [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (578598..583995)	-	5398 b
14	LOC103790407 ID: 103790407	serine protease 55-like [<i>Callithrix jacchus</i> (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (400009..406777)	-	6769 b

15	LOC103790406 ID: 103790406	uncharacterized LOC103790406 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (299285..303830, complement)		4546 b
16	LOC100895146 ID: 100895146	uncharacterized LOC100895146 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (622249..642512, complement)		20264 b
17	LOC100894600 ID: 100894600	uncharacterized LOC100894600 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (572217..577788)		5572 b
18	IL9R ID: 100408034	interleukin 9 receptor [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2819940..2832774)		12835 b
19	VAMP7 ID: 100407430	vesicle associated membrane protein 7 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2745938..2806538)		60601 b
20	LOC100407071 ID: 100407071	ubiquitin-conjugating enzyme E2 Q2-like [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2627627..2674680, complement)	UBE2Q2	47054 b
21	RCN2 ID: 100406707	reticulocalbin 2 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1616984..1635453, complement)		18470 b
22	PSTPIP1 ID: 100406340	proline-serine-threonine phosphatase interacting protein 1 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1520536..1569874, complement)		49339 b
23	PEAK1 ID: 100405972	pseudopodium enriched atypical kinase 1 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1136460..1448598)		312139 b
24	LINGO1 ID: 100405601	leucine rich repeat and Ig domain containing 1 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (885888..908415)		22528 b
25	TMEM266 ID: 100404877	transmembrane protein 266 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2374616..2525645, complement)	CYH15orf27	151030 b
26	ETFA ID: 100404521	electron transfer flavoprotein alpha subunit [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2269977..2362847)		92871 b
27	ISL2 ID: 100404151	ISL LIM homeobox 2 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (2229256..2236409, complement)		7154 b
28	SCAPER ID: 100402952	S-phase cyclin A associated protein in the ER [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1671692..2223259)		551568 b

29	TSPAN3 ID: 100401992	tetraspanin 3 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1487141..1512034)		24894 b
30	HMG20A ID: 100401619	high mobility group 20A [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (1053076..1136219, complement)		83144 b
31	LOC100401245 ID: 100401245	endonuclease 8-like 1 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (610439..617967)		7529 b
32	LOC100400877 ID: 100400877	COMM domain-containing protein 4 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (597837..602906)	COMMD4	5070 b
33	RBMY ID: 100400516	RNA-binding motif protein, Y chromosome, family 1 member B-like [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (520700..537537, complement)		16838 b
34	TSPY ID: 100400155	testis-specific Y-encoded protein 3 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (410732..413800)		3069 b
35	EIF1AY ID: 100399792	eukaryotic translation initiation factor 1A, Y-linked [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (360138..395149)		35012 b
36	LOC100399067 ID: 100399067	chromodomain Y-like protein pseudogene [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (246003..247128, complement)		1126 b
37	LOC100398707 ID: 100398707	THO complex subunit 4 pseudogene [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (195167..209579, complement)		14413 b
38	LOC100398345 ID: 100398345	testis-specific chromodomain protein Y 1-like [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (191011..192826, complement)		1816 b
39	LOC100397988 ID: 100397988	G patch domain-containing protein 2-like [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (126372..156291)		29920 b
40	LOC100397626 ID: 100397626	calcium binding protein 39 pseudogene [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (169098..171500, complement)		2403 b
41	LOC100397261 ID: 100397261	putative GTP-binding protein 6 [Callithrix jacchus (white-tufted-ear marmoset)]	Chromosome Y, NC_013919.1 (14168..21954, complement)	GTPBP6	7787 b