

Índice

1. RESUMEN	10
2. ABSTRACT	12
3. MARCO TEÓRICO	13
3.1. Proyectos genómicos	13
3.2. Visualizadores de contexto local	16
3.3. Visualizadores de contexto global	19
3.4. Lenguaje y entorno de trabajo R	22
4. PLANTEAMIENTO DEL PROBLEMA	23
5. OBJETIVOS	24
5.1. Objetivo General	24
5.2. Objetivos específicos	24
6. MATERIALES Y MÉTODOS	25
6.1. Ingeniería de software	25
6.1.1. Funciones del sistema	25
6.1.2. Atributos	31

6.1.3.	Atributos por función	31
6.1.4.	Actores	35
6.1.5.	Casos de uso	36
6.1.6.	Diagrama de secuencia o de colaboración	54
6.1.7.	Diagrama general de casos de uso	59
6.1.8.	Diagrama de arquitectura del software	60
6.2.	Recopilación y depuración del código para graficar datos en contexto cromosómico	61
6.2.1.	Script generador de gráficos en contexto cromosómico	61
6.3.	Implementación de nuevas funcionalidades	63
6.3.1.	Doble categorización de elementos genómicos	64
6.3.2.	Interoperabilidad con el paquete GenomeGraphs	64
6.3.3.	Elementos genómicos en el cuerpo del cromosoma	64
6.3.4.	Sintenia	65
6.3.5.	BiomaRt	67
6.4.	Empaquetamiento y documentación del código para facilitar su utilización	68
6.4.1.	Creación de un paquete para el ambiente de R	68
6.4.2.	Creación de paquetes de datos	68

6.5. Creación vignette (tutorial) para mostrar el modo de uso del paquete creado	70
7. RESULTADOS	71
7.1. Recopilación y depuración del código para graficar datos en contexto cromosómico	74
7.2. Implementación de nuevas funcionalidades	74
7.2.1. Doble categorización de elementos genómicos	74
7.2.2. Interoperabilidad con paquete GenomeGraphs	76
7.2.3. Elementos genómicos en el cuerpo del cromosoma	78
7.2.4. Sintenia	79
7.2.5. BiomaRt	81
7.3. Empaquetamiento y documentación del código para facilitar su utilización	82
7.3.1. Paquete de la función	82
7.3.2. Paquete de datos	83
7.4. Creación de un vignette (tutorial) para demostrar el modo de uso del paquete	88
7.5. Implementación	89
7.5.1. Creación de histograma	89
7.5.2. Obtención de datos de anotación actualizados	93

8. DISCUSIÓN	97
8.1. Doble categorización de elementos genómicos	100
8.2. Interoperabilidad con el paquete GenomeGraphs	100
8.3. Elementos genómicos en el cuerpo de cromosomas	101
8.4. Sintenia	101
8.5. Trabajo futuro	105
9. CONCLUSIONES	105
10. REFERENCIAS	106
11. ANEXO	110
11.1. Detalle de programas	110
11.2. Tabla comparativa resumen	126
11.3. Tutorial	128
11.4. Código chromPlot	153

Índice de cuadros

1.	Requerimientos del sistema.	30
2.	Descripción de los atributos descritos por el FURPS.	31
3.	Funciones con sus correspondientes atributos FURPS.	34
4.	Descripción de los actores presentes en el programa.	35
5.	Listado de casos de usos.	36
6.	Diagramas de secuencia.	58
7.	Descripción de la línea resumen perteneciente al formato axt.	66
8.	Campos del archivo <i>hg19_cytoBandIdeo</i> para humano.	84
9.	Campos del archivo <i>hg19_gap</i> para el humano.	84
10.	Campos del archivo <i>mm10_cytoBandIdeo</i> para el ratón.	86
11.	Campos del archivo <i>mm10_gap</i> para el ratón.	86
12.	Caso de uso en formato extendido.	90
13.	Flujo de eventos para obtener información de anotación actualizada. . . .	95
14.	Listado de características o funcionalidades encontradas en los programas revisados y en <i>chromPlot</i>	99

Índice de figuras

1.	Flujo de trabajo en proyectos genómicos	15
2.	UCSC Genome Browser	17
3.	Visualización de fenotipos en el genoma completo por medio de Pheno- Gram	21
4.	Diagrama general de casos de usos planteados.	59
5.	Diagrama de arquitectura del software: representación de 3 capas: Pre- sentación, core y persistencia.	60
6.	Flujo de trabajo para la exploración y depuración del script creado en el laboratorio de GENOMED	62
7.	Creación de paquetes de datos	69
8.	Ancestría local en cromosomas humanos.	71
9.	Diversidad genética entre la población chilena y europea.	72
10.	Visualización de expresión diferencial en los cromosomas 15 al 19 del ratón.	73
11.	Categorización por fenotipo y ancestría.	75
12.	Representación de elementos genómicos en una región específica del cromosoma 3.	77
13.	Densidad de SNPs en cuerpo del cromosoma.	78

14. Sintenía entre humano y ratón	80
15. Densidad génica utilizando BiomaRt.	81
16. Diagrama del caso de uso 4.	90
17. Histograma resultante de la función chromhist().	92
18. Diagrama del caso de uso 3.	94
19. Representación del funcionamiento de chromPlot.	104
20. Posibles representaciones dependiendo de las opciones del usuario y las características de los datos.	130
21. Los objetos <code>filtxs</code> y <code>seltxs</code> son subsets de <code>array</code> con datos filtrados por detección del gene (verde) y seleccionados por expresión diferencial (azul). Dado que <code>filtxs</code> y <code>seltxs</code> , contienen información filtrada y seleccionada respectivamente, el histograma resultante es bastante pequeño en comparación con el que representa la densidad génica (rojo).	138
22. Densidad de polimorfismos de único nucleótido (SNP) representada en el cuerpo cromosómico.	143
23. sintenía entre humano y ratón dentro del cromosoma. <code>chromplot</code> puede leer archivos de alineamiento en formato AXT.	144
24. Figura usada en Verdugo et al 2010 (<i>BMC Biology</i> 8:96) para representar asociación entre regiones de divergencia genética (regiones marrón en el cuerpo del cromosoma), QTLs (barras de colores en el lado izquierdo del cromosoma, y ausencia de asociación con densidad génica (histograma al lado derecho del cromosoma).	150