

Creating an history: All analysis will be found in this history

The screenshot displays the Galaxy web interface. At the top, a navigation bar includes 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. A notification banner at the top center states '0 datasets have been deleted permanently' with a green checkmark icon. On the left, a 'Tools' sidebar lists various categories such as 'Get Data', 'Send Data', 'BASIC TOOLS', 'SEQUENCE ANALYSIS', and 'NGS ANALYSIS'. On the right, a 'History' dropdown menu is open, showing options like 'Create New', 'Copy History', 'Share or Publish', 'Delete', and 'Delete Permanently'. The main workspace is currently empty.

Renaming the History (easier to find analysis)

The screenshot shows the Galaxy web interface. At the top, a green notification bar states "0 datasets have been deleted permanently". The main workspace is empty. On the right, the "History" panel shows a search bar with "VCFHunter" entered and a message: "This history is empty. You can load your own data or get data from an external source". A tooltip "Click to rename history" is visible over the history entry. The left sidebar contains a "Tools" menu with categories like "Get Data", "Send Data", "BASIC TOOLS", "SEQUENCE ANALYSIS", "NGS ANALYSIS", "SNP ANALYSIS", and "VCFtools".

Loading existing dataset: To test tools

Galaxy southgreen

Galaxy

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

Get Data Send Data

BASIC TOOLS

Text Manipulation Filter and Sort Join, Subtract and Group Convert Formats

SEQUENCE ANALYSIS

Gene/Protein prediction EMBOSS Operate on Genomic Intervals Fetch Sequences Genomics Fetch Alignments Extract Features

NGS ANALYSIS

NGS: Quality Control NGS: Mapping NGS: GATK Tools NGS: GATK2 Tools NGS: SAM/BAM Manipulations NGS: RNASeq NGS: Assembly NGS: Small RNAs Bedtools Picard Tools

SNP ANALYSIS

NGS: SNP Calling Varscan Population structure GWAS VCFtools Tassel GBS (Version 4.0) Rice Variant Analysis (Rice 3k, IRIGIN, High Density Rice Array (HDRA, 700k SNPs))

galaxy.southgreen.fr/galaxy/library/list

Search Galaxy southgreen

Data Libraries

- Histories
- Workflows
- Visualizations
- Pages

South Green bioinformatics platform

Welcome to GALAXY

Our pre-configured and validated workflows

NGS analyses SNP calling SNP analysis GWAS

GWAS

SNIPlay3 GWAS workflow: Tassel-based GWAS workflow (GLM model) including population structure and correction for structure (Dereeper et al, 2015)

Input: VCF + Phenotypic tabulated file

Access workflow

Structural variations Chrom. reconstruction Metagenomics Gene families Mosaic Genomes

These workflows as part of the services provided by South Green

How to load big datasets?

1 My file → HPC Cluster → 2 My Galaxy History

Choose FTP file

In order to figure out which tools were made available by our team, please activate the "tool search" functionality from the Options drop-down and type "south green" in the lookup filter.

History

search datasets

VCFHunter

0 b

This history is empty. You can load your own data or get data from an external source

Loading existing dataset: To test tools

The screenshot shows the Galaxy Data Libraries interface. At the top, there is a navigation bar with the Galaxy logo and menu items: Analyze Data, Workflow, Shared Data, Visualization, Admin, Help, and User. A status bar on the right indicates 'Using 19%'. Below the navigation bar, the main content area displays a list of data libraries. The list has columns for 'name!', 'description', and 'synopsis'. A search bar and a 'showing 20 of 36 libraries' indicator are present. A red arrow points to the 'GenomeHarvest' entry in the list. Below the list, a URL is visible: galaxy.southgreen.fr/galaxy/library/list#folders/Feafb646da3b7aac5.

name!	description	synopsis
alexis.dereeper@ird.fr		
anne.dievert@cirad.fr		
Arcad	Illumina raw data	
banque_rna_seq_bff_2015		
BFF	Biomass for the Futur data	
Capture_apple		
Cassava_genome		
fabian.pilet@cirad.fr		
Formation		
gaetan.droc@cirad.fr		
Galaxy4SNIPlay		
Galaxy_trainings_2015		
gautier.sarah@cirad.fr		
GenFam		
GenomeHarvest	GenomeHarvest Project	GenomeHarvest Project
hana.chair@cirad.fr		
juliana@utp.edu.co		
julie.orjuela@adnid.fr		
jux.swt@gmail.com	Cleaning, Mapping	
marilyne.summo@cirad.fr		

Loading existing dataset: To test tools

Galaxy Data Libraries

galaxy.southgreen.fr/galaxy/library/list#folders/Feafb646da3b7aac5

Galaxy southgreen

Using 19%

DATA LIBRARIES showing 5 of 5 items

Libraries / GenomeHarvest

name	description	data type	size	time updated (UTC)
..				
Datasets Test		folder		2017-05-22 08:35 AM
trainings_painting		folder		2018-06-29 03:18 AM
VCFHunter	Chromosome Painting and Genetic linkage	folder		2018-12-05 04:05 PM
WP2		folder		2016-06-16 01:10 PM
WP3		folder		2016-06-16 01:06 PM

showing 5 of 5 items

galaxy.southgreen.fr/galaxy/library/list#folders/F52d6bdfafedbb5e5

Loading existing dataset: To test tools

Galaxy Data Libraries

galaxy.southgreen.fr/galaxy/library/list#folders/F52d6bdfafedbb5e5

Galaxy southgreen

Using 19%

DATA LIBRARIES showing 8 of 8 items

Libraries / GenomeHarvest / VCFHunter

name	description	data type	size	time updated (UTC)
<input type="checkbox"/>				
<input checked="" type="checkbox"/> CartoRef.agp		interval	8.0 KB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> Carto.vcf		vcf	78.7 MB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> chr01_test.vcf		vcf	3.5 MB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> chr02_test.vcf		vcf	3.4 MB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> chr03_test.vcf		vcf	4.4 MB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> chr04_test.vcf		vcf	5.6 MB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> chr09_test.vcf		vcf	4.8 MB	2018-12-05 04:05 PM
<input checked="" type="checkbox"/> Origin.tab		pgm	87 bytes	2018-12-05 04:05 PM

showing 8 of 8 items

Going back to history to analyze data

Galaxy Data Libraries

galaxy.southgreen.fr/galaxy/library/list#folders/F52d6bdfafedbb5e5

Galaxy

Analyze Data

Import selected datasets into history

DATA LIBRARIES

showing 8 of 8 items

Libraries / GenomeHarvest / VCFHunter

<input type="checkbox"/>	name [!]	description	data type	size	time updated (UTC)	
<input checked="" type="checkbox"/>	CartoRef.agp		interval	8.0 KB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	Carto.vcf		vcf	78.7 MB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	chr01_test.vcf		vcf	3.5 MB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	chr02_test.vcf		vcf	3.4 MB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	chr03_test.vcf		vcf	4.4 MB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	chr04_test.vcf		vcf	5.6 MB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	chr09_test.vcf		vcf	4.8 MB	2018-12-05 04:05 PM	
<input checked="" type="checkbox"/>	Origin.tab		pgm	87 bytes	2018-12-05 04:05 PM	

showing 8 of 8 items

Going back to history

The screenshot shows the Galaxy web interface. The main content area displays the South Green bioinformatics platform logo and a 'Welcome to GALAXY' message. Below this, there are sections for 'Our pre-configured and validated workflows' and 'Chromosome reconstruction'. A diagram illustrates the workflow: 'My file' (1) is processed by an 'HPC Cluster' (2) and then loaded into 'My Galaxy History'. A note below the diagram states: 'In order to figure out which tools were made available by our team, please activate the "tool search" functionality from the Options drop-down and type "south green" in the lookup filter.'

On the right side, the 'History' panel is visible, showing a list of datasets. A red circle highlights the following datasets:

- VCFHunter (8 shown, 100.33 MB)
- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.aqp

Loaded datasets

Chromosome painting with vcfHunter tool

Developed to answer banana problematics



Musa balbisiana
 $2n = 2x = 22$
 1 to 1.2pg
B genome



Musa acuminata
 $2n = 2x = 22$
 1.1 to 1.3pg
A genome



Nearly half of the banana production worldwide rely on interspecific hybrids of various ploidy (**AB**, **AAB**, **ABB**, **AAAB**).

What is the composition of A and B genomes along chromosomes of cultivated banana hybrids?

Chromosome painting with vcfHunter tool

What is the contribution of ancestral genomes along chromosomes of cultivated hybrids?

Several tools developed under vcfHunter toolbox for this purpose:

The screenshot shows the Galaxy web interface. On the left, the 'Tools' sidebar lists various tools, with 'vcfHunter' circled in red and labeled with a red '2'. Below it, 'vcf2allPropAndCov' is also circled in red and labeled with a red '1'. A red arrow points from the 'vcfHunter' tool towards the main content area. The main content area features the 'South Green bioinformatics platform' logo and a 'Welcome to GALAXY' message. Below this, there is a section titled 'Our pre-configured and validated workflows' with a 'SNP calling' workflow highlighted. The workflow description states: 'The SNP Calling is based on the GATK toolkit, using either UnifiedGenotyper or HaplotypeCaller module. Input: BAM alignment files + FASTA for reference. Output: VCF (Variant call Format) file.' To the right of the workflow description is a vertical list of tool categories: SNP analysis, GWAS, Structural variations, Chrom. reconstruction, Metagenomics, Gene families, and Mosaic Genomes. Below the workflow description is a diagram titled 'How to load big datasets?' showing a flow from 'My file' to 'HPC Cluster' (with a red 'F' icon) and then to 'My Galaxy History'. A 'Choose FTP file' button is located below the HPC Cluster. A text box at the bottom of the diagram reads: 'In order to figure out which tools were made available by our team, please activate the "tool search" functionality from the Options drop-down and type "south green" in the lookup filter.'

Chromosome painting with vcfHunter tool

Used data: A file containing ancestral (non admixed) accession origin

The screenshot shows the Galaxy web interface. On the left, a list of tools is visible, including 'vcfHunter'. The main panel displays a list of accessions and their ancestral origins:

P2	AA
T01	BB
T02	BB
T03	AA
T04	AA
T05	AA
T06	AA
T07	AA
T08	BB
T10	AA
T11	AA

Annotations with arrows point to the 'AA' column, labeled 'Ancestral origin', and the 'T11' row, labeled 'Accession name in the vcf'. The right sidebar shows a history of datasets, with '8: Origin.tab' highlighted and circled in red.

Chromosome painting with vcfHunter tool

Used data: Several vcf files containing the genotypes of 15 accessions

The screenshot shows the Galaxy web interface. The main panel displays the output of the vcfHunter tool, showing a list of contigs and a table of variant calls. The right-hand 'History' panel shows a list of datasets, with five vcf files highlighted in green and circled in red. Red arrows and numbers 1, 2, and 3 point to these files, the vcfHunter tool header, and the bottom navigation bar respectively.

#CHROM	POS	ID	REF	ALT	QUAL
chr09	26844	.	C	G	.
chr09	38559	.	T	G	.
chr09	38565	.	A	C	.
chr09	38574	.	C	T	.
chr09	38610	.	A	C	.
chr09	38651	.	A	T	.
chr09	38707	.	T	C	.
chr09	38713	.	C	G	.
chr09	42909	.	G	A	.
chr09	47454	.	A	C	.
chr09	51113	.	G	A	.
chr09	58026	.	T	C	.
chr09	63859	.	A	G	.
chr09	63864	.	T	C	.
chr09	63876	.	T	A	.
chr09	63889	.	A	G	.
chr09	63901	.	T	C	.
chr09	63924	.	A	C	.
chr09	67990	.	G	A	.
chr09	68003	.	G	T	.
chr09	68024	.	C	T	.
chr09	73617	.	G	A	.
chr09	78839	.	T	G	.
chr09	78853	.	T	A	.
chr09	78866	.	C	T	.
chr09	79031	.	G	C	.
chr09	79081	.	A	G	.
chr09	80569	.	A	G	.
chr09	80631	.	G	A	.

5 vcf files (1 for each chr) for 15 accessions (11 ancestral and 4 hybrids)

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Admin | Help | User | Using 19%

Tools

- NGS: SNP Calling
- Varscan
- Population structure
- GWAS
- VCFtools
- Tassel GBS (Version 4.0)
- Rice Variant Analysis (Rice 3k)
- IRIGIN, High Density Rice Array (HDRA, 700k SNPs)
- GENOME HARVEST
- TransPo-RG Transfer of Position to Resequenced Genome
- parental SNP - Detect parental SNP of hybrids
- Visualization
- TraceAncestor
- vcfHunter** ← ②
- VCF Filter ← ③
- vcf2allPropAndCov
- vcf2allPropAndCovByChr
- vcf2popNew
- RecombCalculatorDose
- Draw_dot_plot
- KDE_classifier
- METAGENOMICS
- FROGS
- EVOLUTION/PHYLOGENY
- Comparative Genomics
- NCBI BLAST+
- Genfam
- Protein analyses
- STATISTICS/GRAPHICS
- Statistics
- Graph/Display Data
- SOUTHGREEN PROJECTS
- SNIPay3
- GNPAnnot Tools

South Green bioinformatics platform

Welcome to **GALAXY**

Our pre-configured and validated workflows

NGS analyses

- SNP calling

The SNP Calling is based on the GATK toolkit, using either UnifiedGenotyper or HaplotypeCaller module.

Input: BAM alignment files + FASTA for reference
Output: VCF (Variant call Format) file

Access workflow

SNP analysis

- GWAS
- Structural variations
- Chrom. reconstruction
- Metagenomics
- Gene families
- Mosaic Genomes

These workflows as part of the services provided by South Green

How to load big datasets?

① My file → HPC Cluster

② HPC Cluster → My Galaxy History

Choose FTP file

In order to figure out which tools were made available by our team, please activate the "tool search" functionality from the Options drop-down and type "south green" in the lookup filter.

History

search datasets

VCFHunter

8 shown

100.33 MB

- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.aqp

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

vcf2allPropAndCov (Galaxy Version 0.1.0)

vcf collection
No data dataset collection available. 1

origin --origin
8: Origin.tab

accession to work with
A 2 column file containing accession name (col1), origin/group (Col2)

ploidy of the accession
2

ancestral accession can't have missing data
no

all accessions should have the allele
yes

Execute

Author Guillaume Martin
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

vcf2allPropAndCov

Description
This program perform two things based on a vcf.
1. It plots for an accession, the allele coverage along its chromosomes.
2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

Inputs:
VCF collection : one per chromosome or a single vcf for all chromosomes
Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab
P2 AA
T01 BB
T02 BB

Outputs:

History
search datasets 2

VCFHunter
8 shown
100.33 MB

- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.agp

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

The screenshot shows the Galaxy web interface with the **vcf2allPropAndCov** tool configuration. The tool parameters are as follows:

- vcf collection:** No data dataset collection available.
- origin --origin:** 8: Origin.tab
- accession to work with:** (empty field)
- ploidy of the accession:** 2
- ancestral accession can't have missing data:** no
- all accessions should have the allele:** yes

The **Execute** button is highlighted with a red box. Below the tool configuration, the description and inputs are provided:

Description: This program perform two things based on a vcf.
1. It plots for an accession, the allele coverage along its chromosomes.
2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

Inputs:
VCF collection : one per chromosome or a single vcf for all chromosomes
Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab

```
P2 AA
T01 BB
T02 BB
```

The **History** panel on the right shows a list of datasets. A red bracket labeled **1** encompasses the tool options and the History panel. A red arrow labeled **2** points to the **For all selected...** button. A red arrow labeled **3** points to the context menu options for the selected datasets.

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

Create a collection from a list of datasets

Collections of datasets are permanent, ordered lists of datasets that can be passed to tools and workflows in order to have analyses done on each member of the entire group. This... [More help](#)

Start over

chr09_test.vcf	Discard
chr04_test.vcf	Discard
chr03_test.vcf	Discard
chr02_test.vcf	Discard
chr01_test.vcf	Discard

Name:

1 2 → **Create list**

vcf2allPropAndCov

Description

This program perform two things based on a vcf.

1. It plots for an accession, the allele coverage along its chromosomes.
2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

Inputs:

VCF collection : one per chromosome or a single vcf for all chromosomes
Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab

```
P2 AA
T01 BB
T02 BB
```

Outputs:

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

The screenshot shows the Galaxy web interface. The main panel displays the configuration for the **vcf2allPropAndCov** tool (Galaxy Version 0.1.0). The configuration includes:

- vcf collection:** No data dataset collection available.
- origin --origin:** 8: Origin.tab (A 2 column file containing accession name (col1), origin/group (Col2))
- accession to work with:** (empty field)
- ploidy of the accession:** 2
- ancestral accession can't have missing data:** no
- all accessions should have the allele:** yes

The **Execute** button is highlighted with a red circle and arrow labeled '2'. The tool description and inputs/outputs are also visible.

The right-hand **History** panel shows a list of datasets. A red circle and arrow labeled '1' points to the search bar. The history includes:

- 9: VCF.conf (a list of datasets)
- 8: Origin.tab
- 7: chr09_test.vcf (checked)
- 6: chr04_test.vcf (checked)
- 5: chr03_test.vcf (checked)
- 4: chr02_test.vcf (checked)
- 3: chr01_test.vcf (checked)
- 2: Carto.vcf
- 1: CartoRef.app

The left-hand **Tools** panel shows the **vcf2allPropAndCov** tool selected, indicated by a red circle and arrow labeled '2'.

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

Galaxy southgreen.fr/galaxy/

Galaxy Analyze Data Workflow Shared Data Visualization Admin Help User Using 19%

Tools

- Rice Variant Analysis (Rice 3k, IRIGIN, High Density Rice Array (HDRA, 700k SNPs))
- GENOME HARVEST
- TransPo-RG Transfer of Position to Resequenced Genome
- parental SNP - Detect parental SNP of hybrids
- Visualization
- TraceAncestor
- vcfHunter
 - vcf Filter
 - vcf2allPropAndCov
 - vcf2allPropAndCovByChr
 - vcf2popNew
 - RecombCalculatorDDose
 - Draw_dot_plot
- KDE_classifier
- METAGENOMICS
- FROGS
- EVOLUTION/PHYLOGENY
- Comparative Genomics
- NCBI BLAST+
- Genfam
- Protein analyses
- STATISTICS/GRAPHICS
- Statistics
- Graph/Display Data
- SOUTHGREEN PROJECTS
- SNIPPlay3
- GNPAnnot Tools
- GNPAnnot Converters
- ESTik
- Expression data
- SAT
- CMAEE tools

vcf2allPropAndCov (Galaxy Version 0.1.0) Options

vcf collection
9: VCF.conf **1** VCF.conf

origin --origin
8: Origin.tab **2** Origin.tab
A 2 column file containing accession name (col1), origin/group (Col2)

accession to work with
Kunnan **3** Kunnan

ploidy of the accession
2 **4** 2 (diploid)

ancestral accession can't have missing data
no **5** no

all accessions should have the allele
yes **6** yes

Execute **7**

Author Guillaume Martin
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

vcf2allPropAndCov

Description

This program perform two things based on a vcf.

1. It plots for an accession, the allele coverage along its chromosomes.
2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

Inputs:

VCF collection : one per chromosome or a single vcf for all chromosomes
Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab

```
B2 AA
T01 BB
T02 BB
```

Outputs:

History

search datasets

VCFHunter
9 shown
100.33 MB

- 9: VCF.conf (a list of 5 datasets)
- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.aqp

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

The screenshot shows the Galaxy web interface with the **vcf2allPropAndCov** tool configuration. The tool parameters are set as follows:

- vcf collection:** 9: VCF.conf
- origin --origin:** 17: Kunnan_stats.tab
- accession to work with:** (empty)
- ploidy of the accession:** 2
- ancestral accession can't have missing data:** no
- all accessions should have the allele:** yes

The **Execute** button is highlighted. A red callout box with the text "4 files will be generated" points to the history panel on the right. The history panel shows the following outputs:

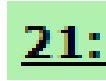
- 17: Kunnan_stats.tab
- 16: Kunnan_AlleleOriginAndRatio.tab
- 15: Kunnan_Ratio.png
- 14: Kunnan_Cov.png
- 9: VCF.conf
- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.agp



Send to cluster



Running on cluster



Finished

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession

The screenshot shows the Galaxy web interface with the **vcf2allPropAndCov** tool configuration. The tool parameters are as follows:

- vcf collection:** 9: VCF.conf
- origin --origin:** 17: Kunnan_stats.tab
- accession to work with:** (empty)
- ploidy of the accession:** 2
- ancestral accession can't have missing data:** no
- all accessions should have the allele:** yes

The **Execute** button is highlighted. Below the configuration, the tool's description and inputs are shown:

vcf2allPropAndCov

Description

This program perform two things based on a vcf.

1. It plots for an accession, the allele coverage along its chromosomes.
2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

Inputs:

- VCF collection : one per chromosome or a single vcf for all chromosomes
- Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab

```
B2 AA
T01 BB
T02 BB
```

Outputs:

The right sidebar shows the **History** panel with a list of datasets:

- 9: VCF.conf
- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.agp

Additional datasets generated by the tool are visible in the history:

- 17: Kunnan_stats.tab
- 16: Kunnan AlleleOriginAndRatio.tab
- 15: Kunnan_Ratio.png
- 14: Kunnan_Cov.png

Chromosome painting with vcfHunter tool

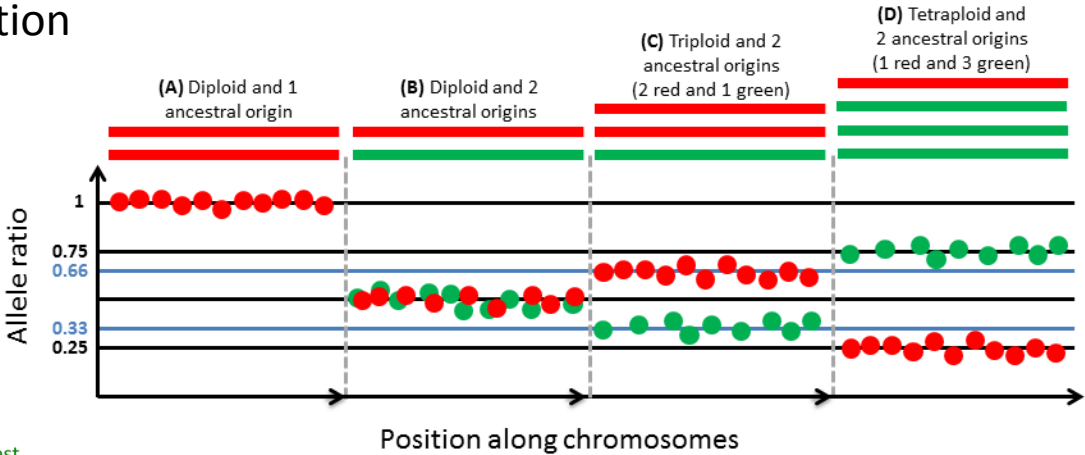
1- Chromosome painting of one accession : How the program work?

- Based on the origin.tab file, attribute an allele to an ancestral group according to the following rule:
 - ✓ If the allele is found **only** in member of an ancestral group (*i.e.* absent from all member of the other(s) group(s)), then the allele is attributed to this group.
- Possible variants to the rule:
 - ✓ Allele should be present in all member of the ancestral group (“yes” or “no”)

all accessions should have the allele

ancestral accession can't have missing data

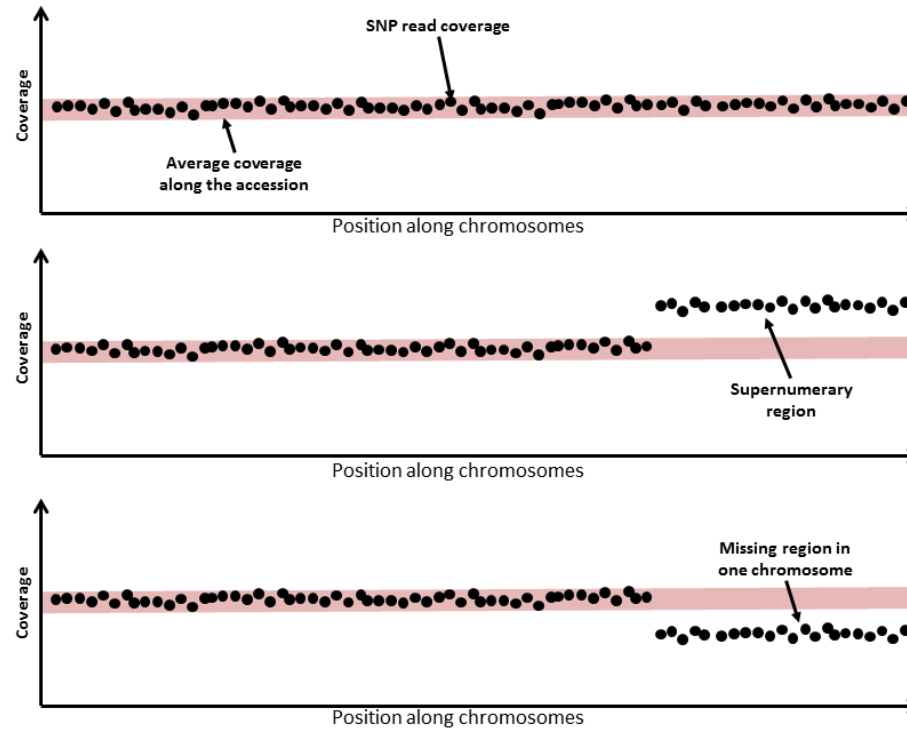
- Calculate, in the studied accession, the proportion of read having this allele
- Plot this proportion



Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : How the program work?

- Plot normalized variant site coverage



Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : Outputs description

Galaxy

Analyze Data Workflow Shared Data Visualization Admin Help User

Tools

- Rice Variant Analysis (Rice 3k, IRIGIN, High Density Rice Array (HDRA, 700k SNPs))
- GENOME HARVEST
- TransPo-RG Transfer of Position to Resequenced Genome
- parental SNP - Detect parental SNP of hybrids
- Visualization
- TraceAncestor
- vcfHunter
- VCF Filter
- vcf2allPropAndCov
- vcf2allPropAndCovByChr
- vcf2popNew
- RecombCalculatorDDose
- Draw_dot_plot
- KDE_classifier
- METAGENOMICS
- FROGS
- EVOLUTION/PHYLOGENY
- Comparative Genomics
- NCBI BLAST+
- Genfam
- Protein analyses
- STATISTICS/GRAPHICS
- Statistics
- Graph/Display Data
- SOUTHGREEN PROJECTS
- SNIPlay3
- GNPAnnot Tools
- GNPAnnot Converters
- ESTik
- Expression data
- SAT
- CMAEE tools

History

search datasets

VCFHunter

13 shown, 8 deleted

103.88 MB

- 21: Kunnan_stats.tab
- 20: Kunnan_AlleleOriginAndRatio.tab
- 19: Kunnan_Ratio.png
- 18: Kunnan_Cov.png
- 9: VCF.conf
- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.aqp

Mean coverage + 1 chromosome

Mean coverage (expected)

Mean coverage - 1 chromosome

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : Outputs description

The screenshot displays the Galaxy web interface for a VCFHunter tutorial. The main workspace shows five chromosome tracks (chr01, chr02, chr03, chr04, chr09) with SNP data points colored in red and green. Below each track is a horizontal bar representing the chromosome, with a red segment and a green segment. The left sidebar lists various tools, including vcfHunter and its sub-tools like VCF Filter and vcf2allPropAndCov. The right sidebar shows the history of datasets, with a red arrow pointing to the file '19: Kunnan_Ratio.png'.

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : Outputs description

The file used to draw the allele ratio figure

1	2	3	4	5
chr02	86818	G	BB	0.22580645161290322
chr02	771450	G	BB	0.5269157894736842
chr02	771450	A	AA	0.47368421052631576
chr02	771450	A	BB	0.55
chr02	771458	G	AA	0.45
chr02	2273733	G	BB	0.5
chr02	2273733	T	AA	0.5
chr02	2273745	T	BB	0.5159515159515151
chr02	2273745	A	AA	0.48484848484848486
chr02	2287602	T	BB	0.44
chr02	2287602	C	AA	0.56
chr02	2343526	C	BB	0.5555555555555556
chr02	2343526	T	AA	0.4444444444444444
chr02	2427206	C	BB	0.5357142857142857
chr02	2427206	G	AA	0.4642857142857143
chr02	2432380	A	BB	0.5952380952380952
chr02	2432380	G	AA	0.40476190476190477
chr02	2432383	G	BB	0.5714285714285714
chr02	2432383	A	AA	0.42857142857142855
chr02	2432393	T	BB	0.5909090909090909
chr02	2432393	C	AA	0.4090909090909091
chr02	2432400	C	BB	0.5348837209302325
chr02	2432400	A	AA	0.46511627906976744
chr02	2447292	T	BB	0.4375
chr02	2447292	A	AA	0.5625
chr02	2447319	T	BB	0.42424242424242425
chr02	2447319	A	AA	0.5757575757575758
chr02	2447328	G	BB	0.375
chr02	2447328	A	AA	0.625
chr02	2480263	G	BB	0.48
chr02	2480263	C	AA	0.52
chr02	2480266	A	BB	0.4782608695652174
chr02	2487990	G	BB	0.4722222222222222
chr02	2487990	A	AA	0.5277777777777778
chr02	2575326	T	BB	0.5714285714285714
chr02	2575326	C	AA	0.42857142857142855
chr02	2696749	T	BB	0.6
chr02	2696749	C	AA	0.4
chr02	4387494	C	BB	0.5
chr02	4387494	T	AA	0.5
chr02	4405457	C	BB	0.5675675675675675
chr02	4405457	G	AA	0.43243243243243246
chr02	4454139	T	BB	0.41379310344827586

Allele ratio in the accession

Allele ancestral group

Attributed allele

Position

Chromosome

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : Outputs description

The screenshot shows the Galaxy web interface with a workflow titled 'VcfHunter/tutorial_Chromosome'. The workflow consists of three steps:

Step	Tool	Input	Output
1	Kunnan	Position_without_missing_data	33426
2	Kunnan	Position_affected_to->AA	1251
3	Kunnan	Position_affected_to->AA:BB	11168

The right-hand side of the interface shows the 'History' panel with a list of datasets. A red arrow points to the dataset '21: Kunnan_stats.tab', which is highlighted in green. Other datasets in the history include '20: Kunnan AlleleOriginAndRatio.tab', '19: Kunnan_Ratio.png', '18: Kunnan_Cov.png', '9: VCF.conf', '8: Origin.tab', '7: chr09_test.vcf', '6: chr04_test.vcf', '5: chr03_test.vcf', '4: chr02_test.vcf', '3: chr01_test.vcf', '2: Carto.vcf', and '1: CartoRef.aqp'.

Allele global statistics

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : To go further

1

The screenshot shows the Galaxy web interface with the **vcf2allPropAndCov** tool selected. The tool configuration panel includes the following fields:

- vcf collection:** 9: VCF.conf
- origin --origin:** 21: Kunnan_stats.tab
- accession to work with:** (empty)
- ploidy of the accession:** 2
- ancestral accession can't have missing data:** no
- all accessions should have the allele:** yes

The description section states: "This program perform two things based on a vcf." and lists two points: 1. It plots for an accession, the allele coverage along its chromosomes. 2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

The **History** panel on the right shows a list of datasets, including **21: Kunnan_stats.tab**, **20: Kunnan AlleleOriginAndRatio.tab**, **19: Kunnan_Ratio.png**, **18: Kunnan_Cov.png**, **9: VCF.conf**, **8: Origin.tab**, **7: chr09_test.vcf**, **6: chr04_test.vcf**, **5: chr03_test.vcf**, **4: chr02_test.vcf**, **3: chr01_test.vcf**, **2: Carto.vcf**, and **1: CartoRef.agp**.

Running the analysis with other options, other accessions of the vcf.
Other hybrids names in the vcf GP1 (triploid), P1 (tetraploid) and P025 (triploid)

Chromosome painting with vcfHunter tool

1- Chromosome painting of one accession : Additional information

1

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Admin | Help | User | Using 19%

Tools

- Rice Variant Analysis (Rice 3k, IRIGIN, High Density Rice Array (HDRA, 700k SNPs))
- GENOME HARVEST
- TransPo-RG Transfer of Position to Resequenced Genome
- parental SNP - Detect parental SNP of hybrids
- Visualization
- TraceAncestor
- vcfHunter
 - vcf Filter
 - vcf2allPropAndCov**
 - vcf2allPropAndCovByChr
 - vcf2popNew
 - RecombCalculatorDDose
 - Draw_dot_plot
- KDE_classifier
- METAGENOMICS
- FROGS
- EVOLUTION/PHYLOGENY
- Comparative_Genomics
- NCBI_BLAST+
- Genfam
- Protein_analyses
- STATISTICS/GRAPHICS
- Statistics
- Graph/Display_Data
- SOUTHGREEN PROJECTS
- SNIPPlay3
- GNPAnnot_Tools
- GNPAnnot_Converters
- ESThik
- Expression_data
- SAT
- CMAEE_tools

ancestral accession can't have missing data
no

all accessions should have the allele
yes

Execute

Author Guillaume Martin
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

vcf2allPropAndCov

Description

This program perform two things based on a vcf.

- It plots for an accession, the allele coverage along its chromosomes.
- It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site in the accession along chromosomes.

Inputs:

VCF collection : one per chromosome or a single vcf for all chromosomes
Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab

```
P2 AA
T01 BB
T02 BB
```

Outputs:

Cov.png: a png file presenting SNP coverage along the chromosomes.
Ratio.png: a png file presenting ancestral allele proportion at a site along the chromosomes.
AlleleOriginAndRatio.tab: a tabulated file reporting for each sites were an ancestral allele has been attributed, its origin and the proportion of reads supporting this allele. This files contains chromosome (col1), position (col2), allele (col3), ancestral origin (col4) and allele ratio (col5).
stats.tab: a tabulated file reporting various statistics on the alleles of the accession.

For more informations:

<https://github.com/SouthGreenPlatform/VcfHunter>

Citations Show BibTeX

Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemaïque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

History

search datasets

VCFHunter
13 shown, 8 deleted
103.88 MB

- 21: Kunnan_stats.tab
- 20: Kunnan_AlleleOriginAndRatio.tab
- 19: Kunnan_Ratio.png
- 18: Kunnan_Cov.png
- 9: VCF.conf
- 8: Origin.tab
- 7: chr09_test.vcf
- 6: chr04_test.vcf
- 5: chr03_test.vcf
- 4: chr02_test.vcf
- 3: chr01_test.vcf
- 2: Carto.vcf
- 1: CartoRef.aqp

Small description

How to cite

Chromosome painting with vcfHunter tool

2- Comparison of several accessions

1

vcf2allPropAndCovByChr (Galaxy Version 0.1.0)

vcf collection
9: VCF.conf

origin --origin
8: Origin.tab

accessions to work with
T04,T02,Kunnan,GP1,P025,P1

ploidy of accessions
3

ancestral accessions can't have missing data
no

all accessions should have the allele
yes

Execute

Author Guillaume MARTIN : guillaume.martin@cirad.fr, Marion Dupouy : marion.dupouy@cirad.fr, Franc-Christophe Baurens : baurens@cirad.fr

Galaxy integration Aurore Comte

Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

vcf2allPropAndCovByChr

Description

This program perform two things based on a vcf.

1. It plots for a chromosome of all accessions in a vcf, the allele coverage along its chromosomes.
2. It identify, based on known ancestral accessions in the vcf, the alleles specific to each groups and plot the alleles proportion at a site along chromosomes for all accessions.

Inputs:

VCF collection : one per chromosome or a single vcf for all chromosomes

Origin : A 2 column file containing accession name (col1), origin/group (Col2)

Origin.tab

```
P2 AA
T01 BB
T02 BB
```

History

search datasets

VCFHunter
14 shown, 17 deleted, 70 hidden
109.82 MB

91: output

Kunnan AlleleOriginAndRatio.tab

9: VCF.conf
a list of 5 datasets

8: Origin.tab

7: chr09_test.vcf

6: chr04_test.vcf

5: chr03_test.vcf

4: chr02_test.vcf

3: chr01_test.vcf

2: Carto.vcf

1: CartoRef.aqp

- 1 VCF.conf
- 2 Origin.tab
- 3 T04,T02,Kunnan,GP1,P025,P1
- 4 3 (triploid -not very important-)
- 5 no
- 6 yes

Chromosome painting with vcfHunter tool

2- Comparison of several accessions: output description

The screenshot shows the Galaxy web interface. At the top, there are browser tabs for 'VcfHunter/tutorial_Chromos...', 'Stockholm format - Wikipedia', and a search bar containing 'stockholm format'. The main interface has a 'Tools' sidebar on the left with categories like 'GENOME HARVEST', 'KDE_classifier', 'METAGENOMICS', etc. The central workspace contains a green notification box with a checkmark icon and the text: '1 job has been successfully added to the queue - resulting in the following datasets: You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.' To the right, the 'History' panel shows a list of datasets. A red arrow points to the entry '102: output' which is highlighted in green. Below it, a list of files is shown, including '21: Kunnan_stats.tab', '20: Kunnan_AlleleOriginAndRatio.tab', '19: Kunnan_Ratio.png', '18: Kunnan_Cov.png', '9: VCF.conf', '8: Origin.tab', '7: chr09_test.vcf', '6: chr04_test.vcf', '5: chr03_test.vcf', '4: chr02_test.vcf', '3: chr01_test.vcf', '2: Carto.vcf', and '1: CartoRef.aqp'. A red rounded rectangle with white text is overlaid on the right side of the interface, containing the text 'A collection of several files'.

Chromosome painting with vcfHunter tool

2- Comparison of several accessions: output description

1 job has been successfully added to the queue - resulting in the following datasets:
You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

**A collection of several files
For each chromosomes:**

- A coverage file
- A ratio file

The number of files per chromosome depends on the of the number of accessions treated (there can be at most 15 accessions per picture)

Chromosome painting with vcfHunter tool

2- Comparison of several accessions: output description

Galaxy interface showing chromosome painting results for several accessions (T04, T02, Kunnant, GP1, P025, P1). The P025 track is highlighted with a blue box and labeled "Missing one haplotype!". A red box highlights a region in the P025 track labeled "Expected ploidy level".

Warning: Because it is a mix of accession with distinct ploidy level, the green threshold (+1 or -1) are only accurate for triploid accessions.

Chromosome painting with vcfHunter tool

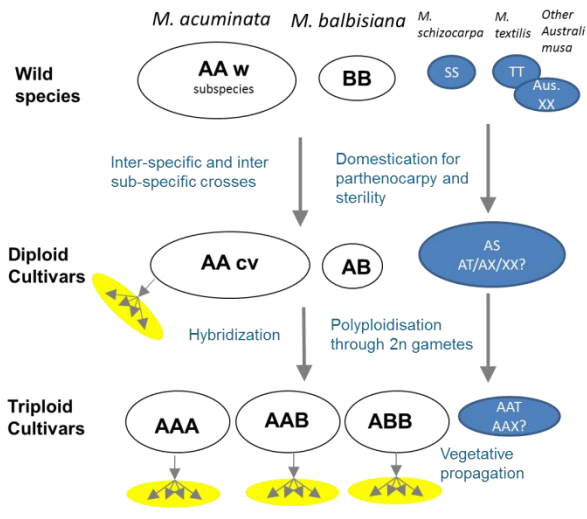
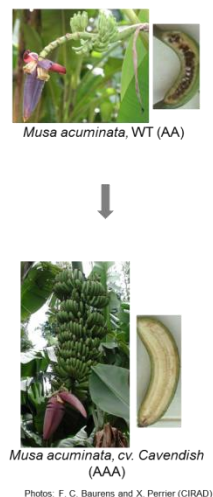
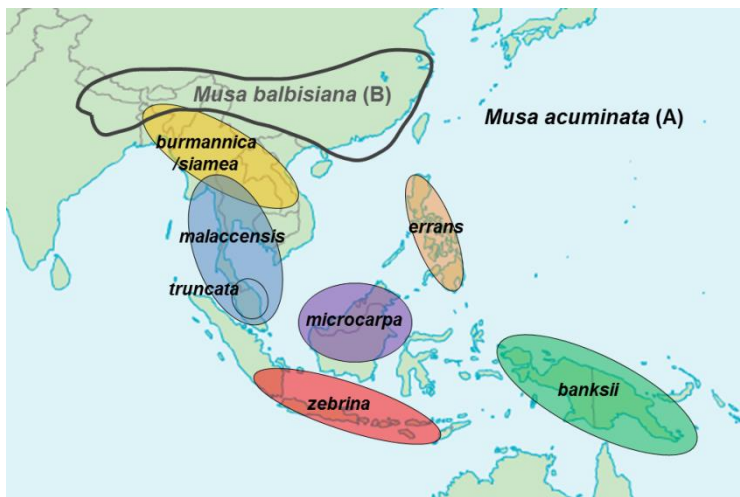
2- Comparison of several accessions: output description

Galaxy interface showing chromosome painting results for six accessions: T04, T02, Kunnann, GP1, P025, and P1. The tracks show green and red dots representing different genetic origins. A text box with arrows points to the GP1 track, stating "2 doses of green origin" and "1 doses of red origin". The right sidebar shows a list of output files for each chromosome (chr01 to chr09) with "View data" buttons.

Warning: Because it is a mix of accession with distinct ploidy level, the green threshold (+1 or -1) are only accurate for triploid accessions.

Genetic mapping analysis with vcfHunter tool

Developed to answer banana problematics



- Cultivated banana are hybrids between distinct species and subspecies showing chromosomal structural rearrangements

What are the chromosomal structures of species and subspecies implicated in cultivated bananas?

What are the consequences of these structural variations on chromosomal pairing, recombination and segregation?

Genetic mapping analysis with vcfHunter tool

What are the chromosomal structures of plants?
What are the consequences of these structural variations on chromosomal pairing, recombination and segregation

Several tools developed under vcfHunter toolbox for this purpose:

The screenshot shows the Galaxy web interface. On the left sidebar, the 'Tools' section is expanded to 'vcfHunter', with a red arrow labeled '2' pointing to it. Below it, 'RecombCalculatorDDose' is highlighted with a red box and a red arrow labeled '1' pointing to it. The main content area displays the 'South Green bioinformatics platform' logo and a 'Welcome to GALAXY' message. Below this, there is a section titled 'Our pre-configured and validated workflows' with a 'SNP calling' workflow highlighted. A diagram titled 'How to load big datasets?' shows a workflow: 'My file' (with a red 'E' icon) is processed by an 'HPC Cluster' (with a red 'E' icon) and then uploaded to 'My Galaxy History'. A red arrow labeled '1' points to the 'My file' step, and a red arrow labeled '2' points to the 'My Galaxy History' step. A text box below the diagram explains: 'In order to figure out which tools were made available by our team, please activate the "tool search" functionality from the Options drop-down and type "south green" in the lookup filter.'

Genetic mapping analysis with vcfHunter tool

Used data: A vcf file containing genotypes for each individuals of a mapping population

The screenshot shows the Galaxy web interface. At the top, there are browser tabs for 'Accueil - CIRAD', 'Galaxy', and 'Guadeloupe-2019 - Google Drive'. The address bar shows 'galaxy.southgreen.fr/galaxy/'. The main interface has a top navigation bar with 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. Below this is a 'Tools' sidebar on the left with various tool categories like 'Get Data', 'BASIC TOOLS', 'SEQUENCE ANALYSIS', 'EMBOSS', 'NGS ANALYSIS', 'SNP ANALYSIS', and 'VCFtools'. The central panel displays a VCF file header and a table of genomic data. The right sidebar shows a 'History' panel with a list of datasets, including 'VCFHunter', '102: output', '21: Kunnan_stats.tab', '20: Kunnan AlleleOriginAndRatio.tab', '19: Kunnan_Ratio.png', '18: Kunnan_Cov.png', '9: VCF.conf', '8: Origin.tab', '7: chr09_test.vcf', '6: chr04_test.vcf', '5: chr03_test.vcf', '4: chr02_test.vcf', '3: chr01_test.vcf', '2: Carto.vcf', and '1: CartoRef.app'. A red circle highlights the '4: chr02_test.vcf' dataset, with a red arrow pointing to it.

```
##fileformat=VCFv4.2
##FILTER=<ID=SnpcCluster,Description="SNPs found in clusters">
##FORMAT=<ID=AD,Number=.,Type=Integer,Description="Allelic depths for the ref and alt alleles in the order listed">
##FORMAT=<ID=DP,Number=1,Type=Integer,Description="Read Depth">
##FORMAT=<ID=GC,Number=1,Type=Float,Description="Ratio between best genotype probability and second best genotype probability">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##GATKCommandLine.VariantFiltration=<ID=VariantFiltration,CommandLineOptions="analysis_type=VariantFiltration input_file=[] showFullBamList=false read_buffer_size=
##Additional.filter=<ID=TAGRemoval,Description="VariantSnpCluster are removed",Date="2018-01-08 14:37:19.913025">
##Additional.filter=<ID=CoverageFiltration,Description="Genotype having less than 15 x coverage and less than 3 x coverage for each allele are converted to missin
##Additional.filter=<ID=MissingDataFiltration,Description="SNP with more than 11 genotype missing are removed",Date="2018-11-14 13:46:44.848676">
##contig=<ID=chr01,length=29070452>
##contig=<ID=chr02,length=29511734>
##contig=<ID=chr03,length=35020413>
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT P134 P135 P136 P137 P039 P038 P132 P133 P035 P034 P037
chr01 5163 . A C . PASS . GT:AD:DP:GC 0/0/0:58,0:58:3.26470499283e+18 0/0/1:71,11:82:5.03108366703e+12 0/0/1:53,7:
chr01 5167 . T C . PASS . GT:AD:DP:GC 0/0/0:57,0:57:1.47610472476e+18 0/0/1:71,11:82:5.03108366703e+12 0/0/1:54,7:
chr01 8846 . A C . PASS . GT:AD:DP:GC 0/0/0:51,0:51:1.27000083428e+16 0/0/1:56,6:62:193.526687819 0/0/1:43,5:48:1405.:
chr01 8858 . C A . PASS . GT:AD:DP:GC 0/0/1:40,11:51:2.88230376152e+17 0/0/1:50,12:62:1.65030130534e+21 0/0
chr01 8888 . T G . PASS . GT:AD:DP:GC 0/0/1:44,7:51:552522078.414 0/0/0:63,0:63:1.7344819115e+20 0/0/0:45,3:48:4393.:
chr01 8905 . T A . PASS . GT:AD:DP:GC 0/0/0:51,0:51:1.27000083428e+16 0/0/0:59,4:63:74842.0551806 0/0/1:42,5:47:2832.:
chr01 17189 . C A . PASS . GT:AD:DP:GC 0/0/0:46,0:46:2.43891551199e+14 0/0/1:60,21:81:3.02231454904e+23 0/0/1:58,23
chr01 17228 . A T . PASS . GT:AD:DP:GC 0/0/1:27,0:47:5.26129332617e+14 0/0/1:59,21:80:7.55578637259e+22 0/0/1:57,23
chr01 20888 . G A . PASS . GT:AD:DP:GC 0/0/1:39,7:46:13301352470.1 0/0/0:53,0:53:6.07605004837e+16 0/0/0:60,0:60:1.598
chr01 20912 . G A . PASS . GT:AD:DP:GC 0/0/1:27,19:46:65536.0 0/0/0:52,0:52:2.80333245379e+16 0/0/0:59,0:59:7.10277380003
chr01 20916 . T C . PASS . GT:AD:DP:GC 0/0/0:46,0:46:2.43891551199e+14 0/0/1:36,16:52:1.09951162778e+12 0/0/1:44,15
chr01 20934 . G A . PASS . GT:AD:DP:GC 0/0/1:34,12:46:1.75921860444e+13 0/0/0:51,0:51:1.27000083428e+16 0/0/0:59,0:
chr01 21019 . A G . PASS . GT:AD:DP:GC 0/0/1:49,25:74:2.81474976711e+14 0/0/0:53,0:53:6.07605004837e+16 0/0/0:35,0:
chr01 29287 . C A . PASS . GT:AD:DP:GC 0/1/1:19,22:41:64.0 0/0/0:47,1:48:98747997863.5 0/0/1:47,12:59:1.1805916207
chr01 29391 . A T . PASS . GT:AD:DP:GC 0/0/0:41,0:41:4.62632016737e+12 0/0/1:63,9:72:1432921316.62 0/0/1:52,10:62:1.31
chr01 29415 . T G . PASS . GT:AD:DP:GC 0/0/0:41,0:41:4.62632016737e+12 0/0/0:69,0:69:2.05639220187e+22 0/0/0:59,0:59:7.102
chr01 37116 . T C . PASS . GT:AD:DP:GC 0/0/0:45,0:45:1.10765491825e+14 0/0/1:61,12:73:2.17253329946e+18 0/0/1:52,13
chr01 37131 . C T . PASS . GT:AD:DP:GC 0/0/1:33,13:46:1.09951162778e+12 0/1/1:32,41:73:262144.0 0/0/1:47,18:65:2.88
chr01 37155 . C T . PASS . GT:AD:DP:GC 0/0/1:33,13:46:1.09951162778e+12 0/1/1:32,41:73:262144.0 0/0/1:44,18:62:4.50
chr01 50285 . T C . PASS . GT:AD:DP:GC 0/0/0:47,0:47:5.26129332617e+14 0/0/1:69,12:80:2.70381667849e+16 0/0/1:30,12
chr01 50313 . T C . PASS . GT:AD:DP:GC 0/0/0:47,0:47:5.26129332617e+14 0/0/1:67,12:79:5.04415731515e+16 0/0/1:31,12
chr01 50330 . G A . PASS . GT:AD:DP:GC 0/0/1:32,15:47:17179869184.0 0/0/1:41,38:79:64.0 0/0/1:34,9:43:1.1258990684
chr01 53409 . C G . PASS . GT:AD:DP:GC 0/0/0:61,0:61:3.53754895582e+19 0/0/1:41,15:56:4.5035962737e+15 0/0/1:39,11
chr01 53451 . C G . PASS . GT:AD:DP:GC 0/0/1:44,16:60:7.20575940379e+16 0/0/0:56,0:56:6.56046544339e+17 0/0/1:37,14
chr01 59316 . G A . PASS . GT:AD:DP:GC 0/0/1:35,23:58:16777216.0 0/0/0:58,0:58:3.26470499283e+18 0/0/1:55,22:77:37
chr01 59428 . T C . PASS . GT:AD:DP:GC 0/0/1:40,11:51:2.88230376152e+17 0/0/1:47,27:74:1.09951162778e+12 0/0
chr01 67413 . T A . PASS . GT:AD:DP:GC 0/0/0:58,0:58:3.26470499283e+18 0/0/1:61,19:80:1.93428131138e+25 0/0/1:55,17
chr01 67461 . G A . PASS . GT:AD:DP:GC 0/0/0:58,0:58:3.26470499283e+18 0/0/1:60,19:79:4.83570327846e+24 0/0/1:55,17
chr01 67517 . C A . PASS . GT:AD:DP:GC 0/0/1:36,14:50:1.75921860444e+13 0/0/0:58,1:59:4.82026012455e+14 0/0/0:33,8:
chr01 67527 . T G . PASS . GT:AD:DP:GC 0/0/0:50,0:50:5.64444815235e+15 0/0/1:48,11:59:9.29162287927e+18 0/0/1:50,11
chr01 67556 . T C . PASS . GT:AD:DP:GC 0/0/0:50,0:50:5.64444815235e+15 0/0/0:57,2:59:66552899445.0 0/0/0:55,5:60:3.011
chr01 89901 . A C . PASS . GT:AD:DP:GC 0/0/0:52,0:52:2.80333245379e+16 0/0/0:41,0:41:4.62632016737e+12 0/0/0:36,0:36:93382
chr01 89911 . T C . PASS . GT:AD:DP:GC 0/0/0:52,0:52:2.80333245379e+16 0/0/0:41,0:41:4.62632016737e+12 0/0/0:36,0:36:93382
chr01 89917 . C T . PASS . GT:AD:DP:GC 0/0/0:52,0:52:2.80333245379e+16 0/0/1:41,12:53:2.88230376152e+17 0/0/1:37,12
chr01 89923 . G T . PASS . GT:AD:DP:GC 0/0/0:52,0:52:2.80333245379e+16 0/0/1:40,12:52:7.20575940379e+16 0/0/1:37,12
chr01 89958 . T C . PASS . GT:AD:DP:GC 0/0/0:52,0:52:2.80333245379e+16 0/0/1:39,12:51:1.8014385085e+16 0/0/1:37,12
```

Genetic mapping analysis with vcfHunter tool

Used data: An agp file locating scaffolds in the reference assembly

The screenshot shows the Galaxy web interface. The main panel displays a table with 9 columns and multiple rows of genomic data. The table header is "# agp-version 2.0". The data includes chromosome (chr01), coordinates, scaffold names, and various flags. The right-hand panel shows a "History" section with a search bar and a list of datasets. A red circle with the number "1" highlights the dataset "1: CartoRef.agp" in the history list.

1	2	3	4	5	6	7	8	9
# agp-version 2.0								
chr01	1	10853	1	W	scaffold1006	1	10853	+
chr01	10854	10953	2	N	100	fragment	no	
chr01	10954	14478003	3	W	scaffold3	1	14467050	-
chr01	14478004	14478103	4	N	100	fragment	no	
chr01	14478104	15769260	5	W	scaffold58	1	1291157	-
chr01	15769261	15769360	6	N	100	fragment	no	
chr01	15769361	15814576	7	W	scaffold551	1	45216	+
chr01	15814577	15814676	8	N	100	fragment	no	
chr01	15814677	17019194	9	W	scaffold63	1	1204518	-
chr01	17019195	17019294	10	N	100	fragment	no	
chr01	17019295	17819353	11	W	scaffold100	1	800059	+
chr01	17819354	17819453	12	N	100	fragment	no	
chr01	17819454	18738040	13	W	scaffold84	1	918587	+
chr01	18738041	18738140	14	N	100	fragment	no	
chr01	18738141	20132513	15	W	scaffold50	1	1394373	-
chr01	20132514	20132613	16	N	100	fragment	no	
chr01	20132614	20699287	17	W	scaffold152	1	566674	-
chr01	20699288	20699387	18	N	100	fragment	no	
chr01	20699388	20894093	19	W	scaffold290	1	194706	-
chr01	20894094	20894193	20	N	100	fragment	no	
chr01	20894194	21247679	21	W	scaffold216	1	353486	+
chr01	21247680	21247779	22	N	100	fragment	no	
chr01	21247780	22417285	23	W	scaffold65	1	1169506	+
chr01	22417286	22417385	24	N	100	fragment	no	
chr01	22417386	22652210	25	W	scaffold267	1	234825	+
chr01	22652211	22652310	26	N	100	fragment	no	
chr01	22652311	23188813	27	W	scaffold157	1	536503	+
chr01	23188814	23188913	28	N	100	fragment	no	
chr01	23188914	24583804	29	W	scaffold51	1	1394891	-
chr01	24583805	24583904	30	N	100	fragment	no	
chr01	24583905	24963756	31	W	scaffold209	1	379852	-
chr01	24963757	24963856	32	N	100	fragment	no	
chr01	24963857	25777628	33	W	scaffold99	1	813772	+
chr01	25777629	25777728	34	N	100	fragment	no	
chr01	25777729	25974873	35	W	scaffold288	1	197145	+
chr01	25974874	25974973	36	N	100	fragment	no	
chr01	25974974	26399442	37	W	scaffold188	1	424469	+
chr01	26399443	26399542	38	N	100	fragment	no	
chr01	26399543	26446166	39	W	scaffold545	1	46624	+
chr01	26446167	26446266	40	N	100	fragment	no	
chr01	26446267	26479156	41	W	scaffold655	1	32890	-
chr01	26479157	26479256	42	N	100	fragment	no	

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

1

The screenshot displays the Galaxy web interface for the **vcf2popNew** tool (Galaxy Version 0.1.0). The tool is selected in the left-hand 'Tools' panel, indicated by a red circle with the number '1' and an arrow. The main panel shows the tool's configuration options, including the vcf file path, segregation tested options, and various coverage and frequency thresholds. The right-hand panel shows the 'History' tab with a list of output files, including 'RatioAndCov_chr01_1_Cov.png' and 'RatioAndCov_chr01_1_Ratio.png'.

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The screenshot displays the Galaxy web interface for the **vcf2popNew** tool (Galaxy Version 0.1.0). The tool configuration is as follows:

- The vcf file --vcf:** 119: chr09_test.vcf
- Segregation test:** 114: Carto.vcf
- Maximal read coverage for a marker in an accession --MaxCov:** 1000
- Window for minority allele coverage frequency to be insufficient to call a heterozygous but to high to call an homozygous --WinFreq:** 0.01:0.1
- Minimal read number of minor allele to call variant heterozygous --MinAlCov:** 1
- Maximal missing data proportion in the progeny (Excluding parents) --miss:** 0.1
- Add allele coverage information to genotype file --addcov:** no
- accessions to exclude from the filtration --NoUsed:** Nothing selected
- accessions to exclude from the analysis --exclude:** Nothing selected
- The reference fasta file. --ref:** Nothing selected

The **output** panel on the right shows a list of generated files for chromosomes 01, 02, 03, 04, and 09, including 'RatioAndCov' and 'Ratio.png' files.

Carto.vcf

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The screenshot shows the Galaxy web interface with the **vcf2popNew** tool (Galaxy Version 0.1.0) configured. The tool parameters are as follows:

- The vcf file --vcf:** 114: Carto.vcf
- Segregation tested --seg:** (Empty field, highlighted by a red callout box with the text "Marker segregation(s) tested to select markers")
- Minimal read coverage for a marker in an accession --MinCov:** 10
- Maximal read coverage for a marker in an accession --MaxCov:** 1000
- Window for minority allele coverage frequency to be insufficient to call a heterozygous but to high to call an homozygous --WinFreq:** 0.01:0.1
- Minimal read number of minor allele to call variant heterozygous --MinAlCov:** 1
- Maximal missing data proportion in the progeny (Excluding parents) --miss:** 0.1
- Add allele coverage information to genotype file --addcov:** no
- accessions to exclude from the filtration --NoUsed:** Nothing selected
- accessions to exclude from the analysis --exclude:** Nothing selected
- The reference fasta file. --ref:** Nothing selected

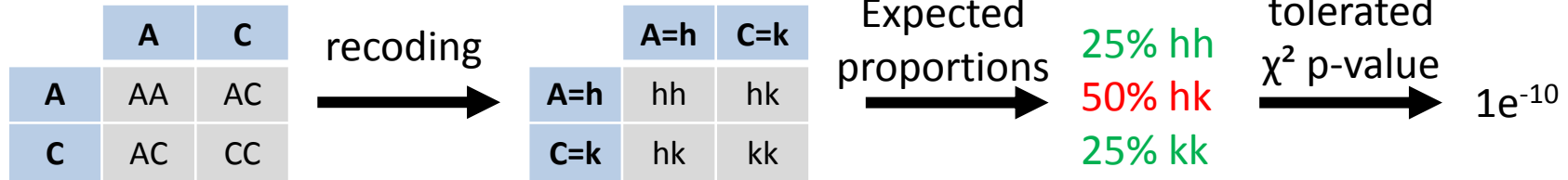
The right-hand side of the interface shows the **History** panel with a list of datasets including **RatioAndCov_chr01_1_Cov.png**, **RatioAndCov_chr01_1_Ratio.png**, **RatioAndCov_chr02_1_Cov.png**, **RatioAndCov_chr02_1_Ratio.png**, **RatioAndCov_chr03_1_Cov.png**, **RatioAndCov_chr03_1_Ratio.png**, **RatioAndCov_chr04_1_Cov.png**, **RatioAndCov_chr04_1_Ratio.png**, **RatioAndCov_chr09_1_Cov.png**, and **RatioAndCov_chr09_1_Ratio.png**.

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The expected marker segregation depends of the cross studied:

- ✓ Selfing of an heterozygous diploid accession:

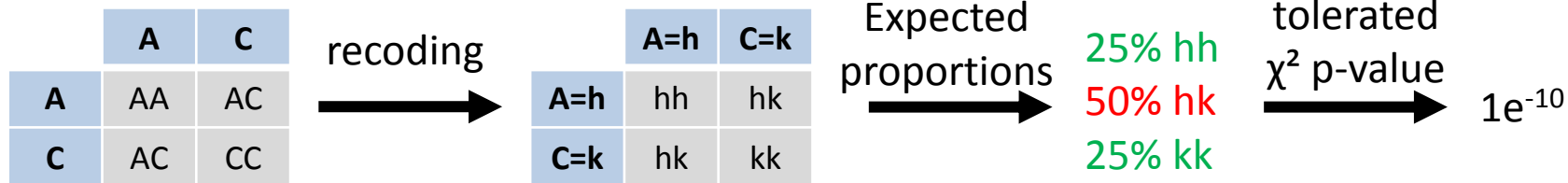


Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The expected marker segregation depends of the cross studied:

- ✓ Selfing of an heterozygous diploid accession:



Homozygous = Ho
Heterozygous = He

Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10

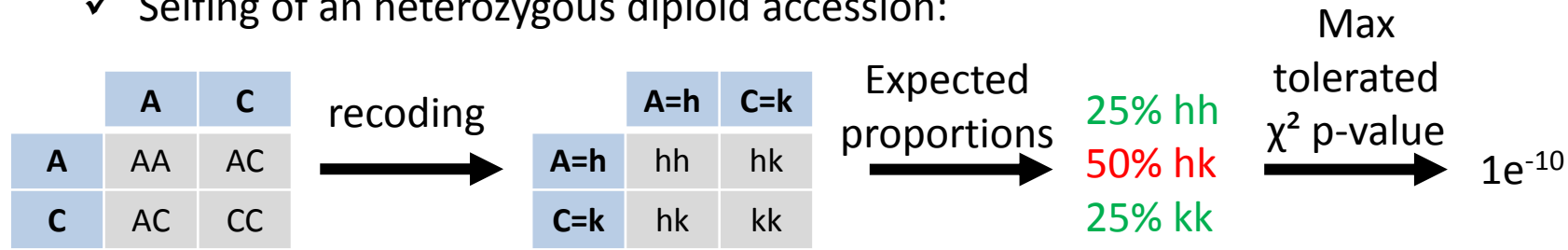
Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The expected marker segregation depends of the cross studied:

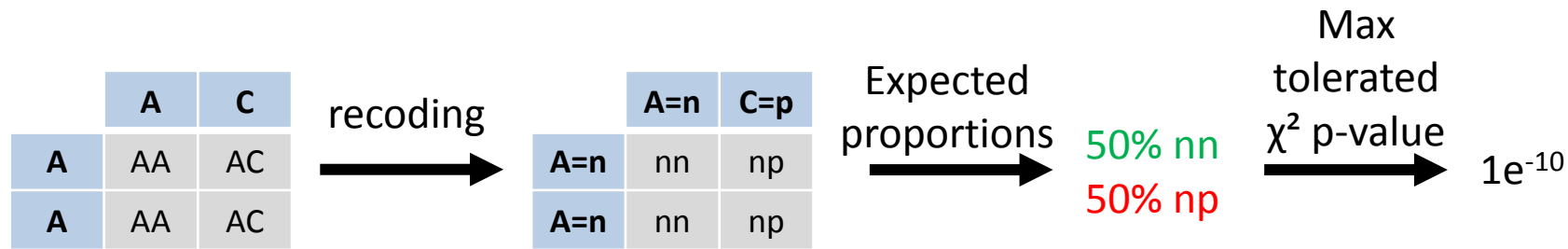
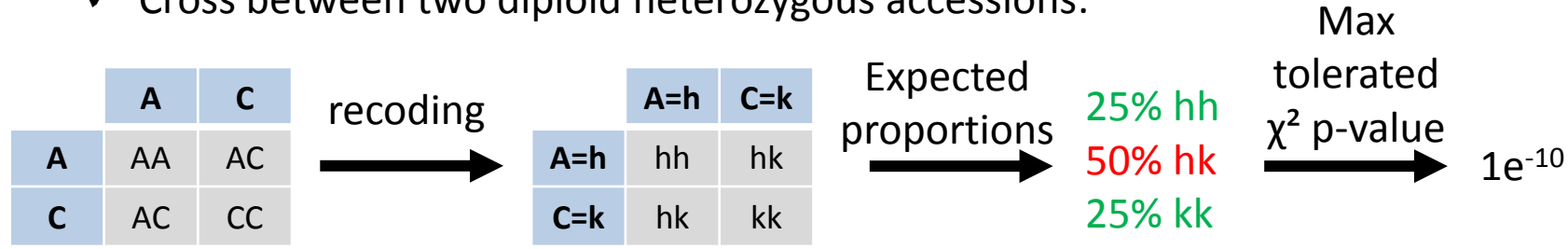
Homozygous = Ho
Heterozygous = He

✓ Selfing of an heterozygous diploid accession:



Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10

✓ Cross between two diploid heterozygous accessions:

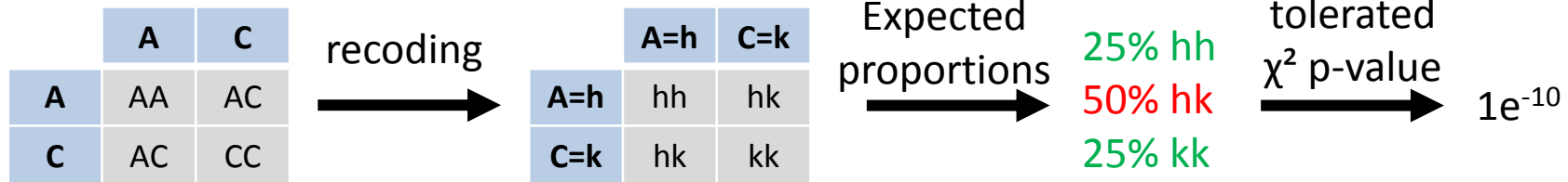


Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The expected marker segregation depends of the cross studied:

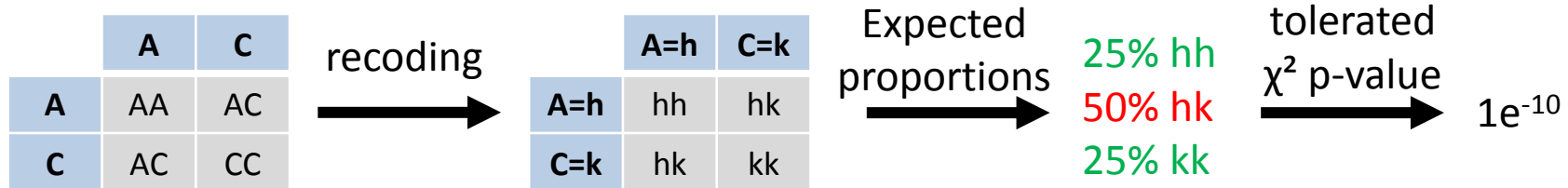
✓ Selfing of an heterozygous diploid accession:



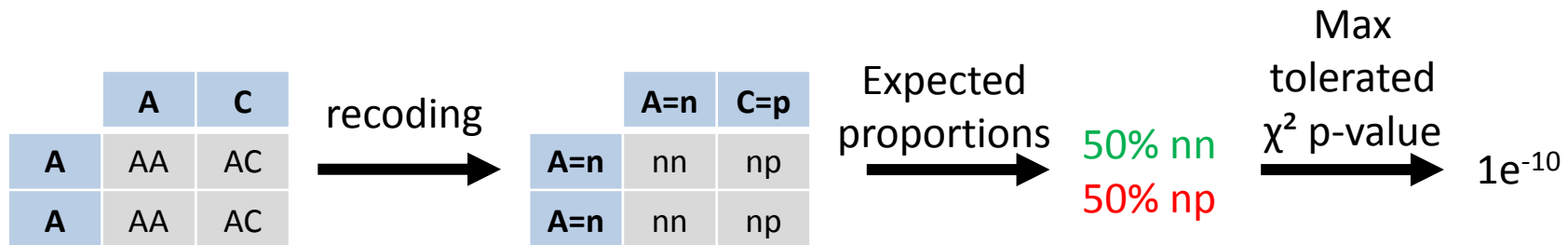
Homozygous = Ho
Heterozygous = He

Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10

✓ Cross between two diploid heterozygous accessions:



Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10



SimpleDose:P1,P2:Ho,He@nn,np:0.5,0.5:1e-10

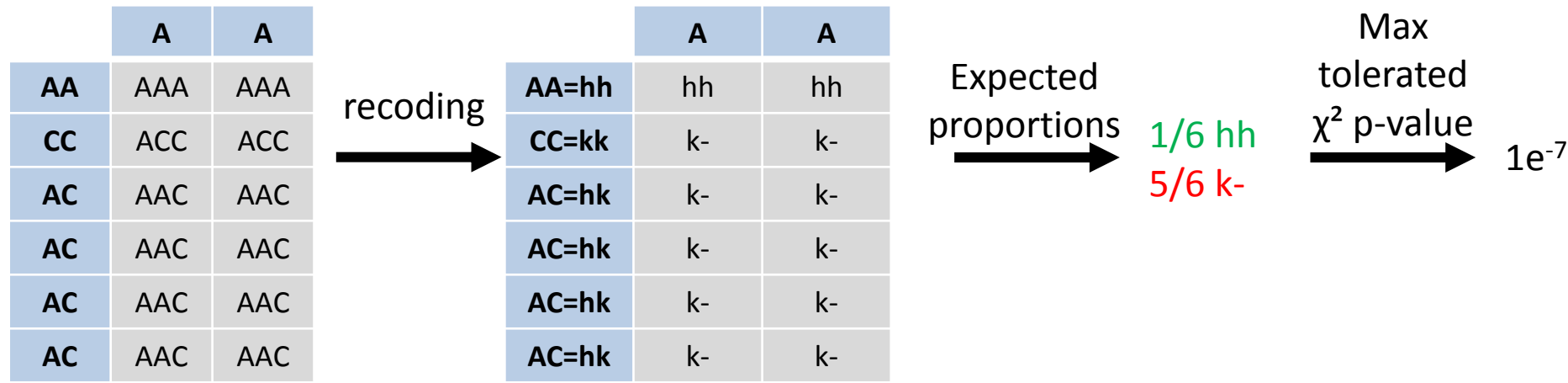
Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

Homozygous = Ho
Heterozygous = He

The expected marker segregation depends of the cross studied:

- ✓ Selfing of an heterozygous diploid accession with an heterozygous tetraploid accession:
An additional segregation tested: double dose markers (e.g: P2 genotype = A/A/C/C, P1 genotype = C/C)



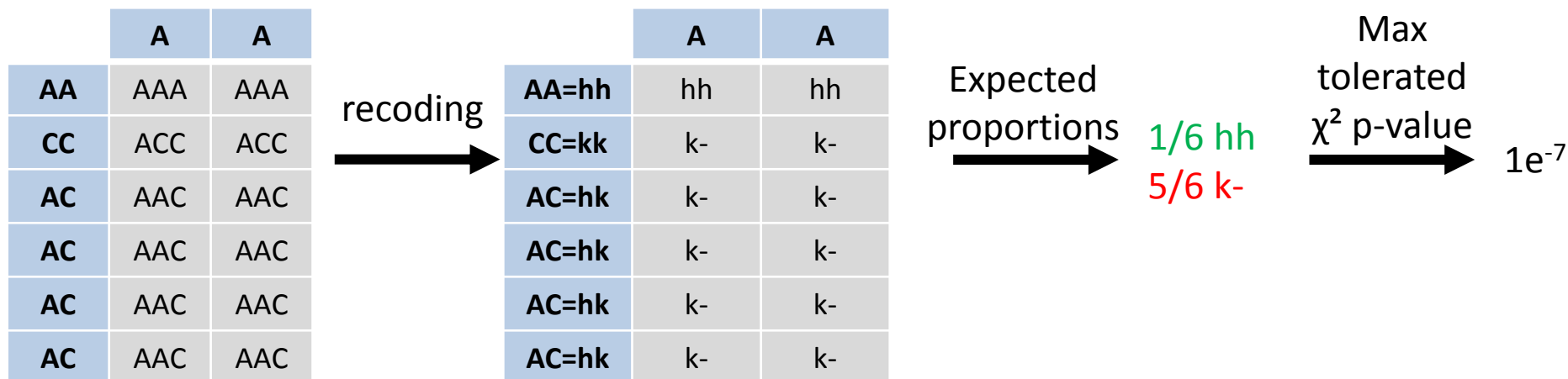
Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

Homozygous = Ho
Heterozygous = He

The expected marker segregation depends of the cross studied:

- ✓ Selfing of an heterozygous diploid accession with an heterozygous tetraploid accession:
An additional segregation tested: double dose markers (e.g: P2 genotype = A/A/C/C, P1 genotype = C/C)



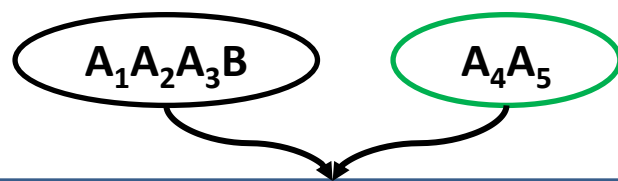
DoubleDose:P1,P2:Ho,He@hh,k-:0.1667,0.8333:1e-7

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The expected marker segregation depends of the cross studied:

- ✓ Selfing of an heterozygous diploid accession with an heterozygous tetraploid accession:
An additional segregation tested: double dose markers
- ✓ This is the type of cross we have in our example!



With some regions which are AABB or $A_1A_1A_2B$ for example

- ✓ We will select markers for several segregations

DoubleDose:P1,P2:Ho,He@hh,k-:0.1667,0.8333:1e-7

Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10

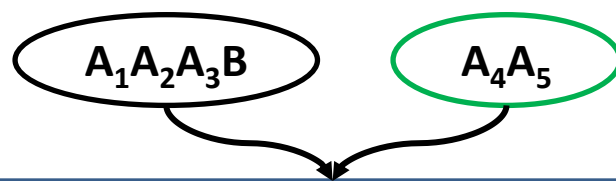
SimpleDose:P1,P2:Ho,He@nn,np:0.5,0.5:1e-10

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The expected marker segregation depends of the cross studied:

- ✓ Selfing of an heterozygous diploid accession with an heterozygous tetraploid accession:
An additional segregation tested: double dose markers
- ✓ This is the type of cross we have in our example!



With some regions which are AABB or $A_1A_1A_2B$ for example

Genotyped progeny (180 individuals)

- ✓ We will select markers for several segregations

DoubleDose:P1,P2:Ho,He@hh,k-:0.1667,0.8333:1e-7

Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10

SimpleDose:P1,P2:Ho,He@nn,np:0.5,0.5:1e-10

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The screenshot displays the Galaxy web interface for the vcf2popNew tool. The tool configuration is as follows:

- The vcf file --vcf:** 114: Carto.vcf
- Segregation tested --seg:** SimpleDose:P1,P2:Ho,He@nn,np:0.5,0.5:1e-10/Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10/DoubleDose:P1,P2:Ho,He@hh,k:-:0.1667,0.8333:1e-7
- Minimal read coverage for a marker in:** 10
- Maximal read coverage for a marker in:** 1000
- Window for minority allele coverage fr:** 0.01:0.1
- Minimal read number of minor allele to call variant heterozygous --MinAlCov:** 1
- Maximal missing data proportion in the progeny (Excluding parents) --miss:** 0.1
- Add allele coverage information to genotype file --addcov:** no
- accessions to exclude from the filtration --NoUsed:** Nothing selected
- accessions to exclude from the analysis --exclude:** Nothing selected
- The reference fasta file. --ref:** Nothing selected

The right-hand panel shows the output files, including:

- RatioAndCov_chr01_1_Cov.png
- RatioAndCov_chr01_1_Ratio.png
- RatioAndCov_chr03_1_Ratio.png
- RatioAndCov_chr04_1_Cov.png
- RatioAndCov_chr04_1_Ratio.png
- RatioAndCov_chr09_1_Cov.png
- RatioAndCov_chr09_1_Ratio.png

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

vcf2popNew (Galaxy Version 0.1.0)

The vcf file --vcf
114: Carto.vcf

Segregation tested --seg
:Ho,He@nn,np:0.5,0.5:1e-10/Bridge:P1,P2:Ho,He,Ho@hh,hk,kk:0.25,0.5,0.25:1e-10/DoubleDose:P1,P2:Ho,He@hh,k:-0.1667,0.8333:1e-7

Minimal read coverage for a marker in an accession --MinCOV
15

Maximal read coverage for a marker in an accession --MaxCOV
300

Window for minority allele coverage frequency to be insufficient to call a heterozygous but to high to call an homozygous --WinFreq
0.01:0.1

Minimal read number of minor allele to call variant heterozygous --MinAlcov
3

Maximal missing data proportion in the progeny (Excluding parents) --miss
0.05

Add allele coverage information to genotype file --addcov
no

accessions to exclude from the filtration --NoUsed
Nothing selected

accessions to exclude from the analysis --exclude
Nothing selected

The reference fasta file. --ref
Nothing selected

1 15
2 300
3 0.01:0.1
4 3
5 0.05
6 no

History

search datasets

VCFHunter
14 shown, 112 deleted, 10 hidden
102.91 MB

126: output
a list of datasets

125: Kunnan_stats.tab

124: Kunnan_AlleleOriginAndRatio.tab

123: Kunnan_Ratio.png

122: Kunnan_Cov.png

121: vcf.conf
a list of 5 datasets

120: Origin.tab

119: chr09_test.vcf

118: chr04_test.vcf

117: chr03_test.vcf

116: chr02_test.vcf

115: chr01_test.vcf

114: Carto.vcf

113: CartoRef.gpp

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: vcf2PopNew

The screenshot shows the Galaxy web interface for the vcf2popNew tool. The left sidebar lists various tool categories such as Tools, GENOME HARVEST, visualization, and vcfHunter. The main content area displays the tool's description, inputs, and outputs. A red circle with the number '2' highlights the 'Execute' button, and another red circle with the number '1' highlights the 'VCF : vcf to work on' input field. The right sidebar shows a history of datasets, including vcfHunter, 126: output, 125: Kunnan_stats.tab, 124: Kunnan_AlleleOriginAndRatio.tab, 123: Kunnan_Ratio.png, 122: Kunnan_Cov.png, 121: vcf.conf, 120: Origin.tab, 119: chr09_test.vcf, 118: chr04_test.vcf, 117: chr03_test.vcf, 116: chr02_test.vcf, 115: chr01_test.vcf, 114: Carto.vcf, and 113: CartoRef.app.

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

The screenshot shows the Galaxy web interface. At the top, a green notification box states: "1 job has been successfully added to the queue - resulting in the following datasets: You can check the status of queued jobs and view the resulting data by refreshing the History pane. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered." Below this, a red callout box contains the text "A collection of several files". The right-hand "History" pane lists 15 datasets, including "137: output", "126: output", "125: Kunnan_stats.tab", "124: Kunnan_AlleleOriginAndRatio.tab", "123: Kunnan_Ratio.png", "122: Kunnan_Cov.png", "121: vcf.conf", "120: Origin.tab", "119: chr09_test.vcf", "118: chr04_test.vcf", "117: chr03_test.vcf", "116: chr02_test.vcf", "115: chr01_test.vcf", "114: Carto.vcf", and "113: CartoRef.app". The left-hand "Tools" pane lists various tools, with "vcfHunter" and its sub-tools like "VCF Filter" and "vcf2allPropAndCov" visible.

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Galaxy interface showing the output of the vcfHunter tool. The main panel displays a table of variant data with columns for chromosome, position, and various population genotypes (P191 to P303). A red box highlights a specific row of data.

2 points to the 'Show all' button in the warning banner at the top.

1 points to the 'Pop.tab' output file in the right-hand 'History' panel.

3 points to a specific row in the variant data table.

Text overlay: List of variant line passing missing data filter (not used after in this tutorial but can be of researcher interest)

chromosome	position	P191	P134	P135	P136	P137	P039	P038	P132	P133	P035	P034	P037	P036	P031	P030	P03
chr01	17228	A/A	A/T	A/T	A/A	A/T	A/A	./.	A/T	A/T	A/A	A/A	A/T	A/A	A/A	A/A	A/T
chr01M20912	20912	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/A	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G
chr01M20916	20916	T/T	T/C	T/C	T/C	T/T	T/T	T/C	T/C	T/C	T/T	T/T	T/C	T/C	T/T	T/T	T/C
chr01M29287	29287	C/A	./.	C/A	C/C	C/A	C/C	C/A	C/A	C/C	C/A	C/C	C/A	C/A	C/A	C/C	C/A
chr01M37116	37116	T/T	T/C	T/C	T/C	T/T	T/T	T/T	T/C	T/C	T/T	T/T	T/C	T/T	T/T	T/T	T/C
chr01M37131	37131	C/T	C/T	C/T	C/T	./.	C/T	C/T	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/T	C/C
chr01M37155	37155	C/T	C/T	C/T	C/T	./.	C/T	C/T	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/T	C/C
chr01M50285	50285	T/T	T/C	T/C	T/C	T/T	T/T	T/T	T/C	T/C	T/T	T/T	T/C	T/T	T/T	T/T	T/C
chr01M50313	50313	T/T	T/C	T/C	T/C	T/T	T/T	T/T	T/C	T/C	T/T	T/T	T/C	T/T	T/T	T/T	T/C
chr01M50330	50330	G/A	G/A	G/A	G/A	G/G	G/A	G/A	G/G	G/A	G/G	G/A	G/G	G/A	G/A	G/A	G/G
chr01M53409	53409	C/G	C/G	C/G	C/G	C/C	C/G	./.	C/C	C/G	C/G	./.	C/C	C/G	C/C	C/C	C/G
chr01M53451	53451	C/G	C/C	C/G	C/C	C/G	C/C	C/G	C/G	./.	C/G	./.	C/G	C/G	C/C	C/C	C/G
chr01M59316	59316	G/A	G/G	G/A	G/G	G/A	G/G	G/A	G/A	G/G	G/A	G/G	G/A	G/A	G/A	G/G	G/A
chr01M59428	59428	T/C	T/C	T/C	T/C	T/T	T/C	T/C	T/T	T/C	T/T	T/C	T/T	T/C	T/T	T/T	T/C
chr01M67527	67527	T/T	T/G	T/G	T/G	./.	T/G	T/T	T/T	T/G	T/G	T/T	T/T	T/G	T/T	T/T	T/G
chr01M89917	89917	C/C	C/T	C/T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/C	C/C	C/T	C/C	C/C
chr01M89923	89923	G/G	G/T	G/T	G/G	G/G	G/G	G/G	G/T	G/T	G/T	G/G	G/T	G/G	G/T	G/G	G/T
chr01M105158	105158	C/T	C/C	C/T	C/C	C/T	C/C	C/T	C/C	C/T	C/C	C/T	C/C	C/T	C/T	C/C	C/C
chr01M127702	127702	T/G	T/T	T/G	T/T	T/T	T/T	T/G	T/G	T/T	T/G	T/T	T/G	T/G	T/T	T/T	T/G
chr01M144740	144740	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
chr01M152926	152926	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
chr01M157399	157399	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
chr01M165977	165977	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
chr01M165990	165990	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
chr01M182847	182847	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A	G/A
chr01M186534	186534	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
chr01M186547	186547	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
chr01M186553	186553	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
chr01M192113	192113	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
chr01M192120	192120	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
chr01M192136	192136	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
chr01M226490	226490	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
chr01M235828	235828	A/A	A/G	A/G	A/G	A/A	A/G	A/A	A/A	A/G	A/G	./.	A/A	A/G	A/A	A/A	A/G
chr01M246911	246911	T/C	T/C	T/C	T/C	T/T	T/C	T/C	T/C	T/T	T/C	T/T	T/T	T/C	T/T	T/C	./.
chr01M251887	251887	C/G	C/G	C/G	C/G	C/C	C/G	./.	C/C	C/G	C/G	C/C	C/C	C/C	C/C	C/C	C/C
chr01M251913	251913	C/C	C/A	C/A	C/A	C/C	C/A	./.	C/C	C/A	C/A	C/C	C/C	C/A	C/C	C/C	C/A
chr01M251923	251923	G/G	G/G	G/G	G/G	./.	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/A
chr01M310320	310320	A/A	A/G	A/G	A/G	A/A	A/G	A/A	A/A	A/G	A/G	A/A	A/A	A/G	A/A	A/A	A/G
chr01M310368	310368	A/A	A/C	A/C	A/A	A/A	A/A	A/A	A/C	A/A	A/C	A/A	A/A	A/A	A/A	A/A	A/A
chr01M326767	326767	G/G	G/T	G/T	G/T	G/G	G/T	G/G	G/A	G/T	G/T	G/G	G/G	G/T	G/G	G/G	G/G
chr01M340656	340656	C/C	C/G	C/G	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/C
chr01M381756	381756	T/A	T/A	T/A	T/A	T/T	T/A	T/T	T/A	T/A	T/A	T/T	T/A	T/T	T/A	T/T	T/A
chr01M389536	389536	C/C	C/T	C/T	C/T	./.	C/T	C/C	C/C	C/T	C/T	C/T	C/C	C/C	C/T	C/C	C/C
chr01M408718	408718	G/A	G/A	G/A	G/A	G/A	G/G	G/G	G/G	G/A	G/A	G/G	G/G	G/A	G/G	G/G	G/A
chr01M415461	415461	C/C	C/T	C/T	C/T	C/C	C/T	C/C	C/C	C/T	C/T	C/T	C/C	C/C	C/T	C/C	C/C
chr01M466101	466101	C/C	C/T	C/T	C/C	C/T	C/C	C/C	C/C	C/T	C/T	C/T	C/C	C/C	C/T	C/C	C/T
chr01M471076	471076	T/T	T/G	T/G	T/G	T/T	T/G	T/T	T/T	T/G	T/G	T/T	T/T	T/T	T/T	T/T	T/T
chr01M471092	471092	G/G	G/T	G/T	G/T	G/G	G/T	G/G	G/T	G/T	G/G	G/T	G/G	G/G	G/T	G/T	G/G
chr01M488027	488027	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

The screenshot displays the Galaxy web interface with the following components:

- Navigation Menu (Left):** Tools, vcfHunter, Tassel GBS (Version 4.0), Rice Variant Analysis (Rice 3k, IRIGIN, High Density Rice Array (HDRA, 700k SNPs)), GENOME HARVEST, TransPo-RG Transfer of Position to Resequenced Genome, parental SNP - Detect parental SNP of hybrids, Visualization, TraceAncestor, vcfHunter, VCF Filter, vcf2allPropAndCov, vcf2allPropAndCovByChr, vcf2popNew, RecombCalculatorDDose, Draw_dot_plot, KDE_classifier, METAGENOMICS, FROGS, EVOLUTION/PHYLOGENY, Comparative Genomics, NCBI BLAST+, Genfam, Protein analyses, STATISTICS/GRAPHICS, Statistics, Graph/Display Data, SOUTHGREEN PROJECTS, SNPPlay3, GNPAnnot Tools, GNPAnnot Converters, ESTHik, Expression data, SAT.
- Main Data Table:** A table with columns for marker IDs (P148-P118, P118), genotypes (e.g., A/A, G/A, T/C), and statistical values (P-value, ChiSquare). Two red callout boxes highlight the 'X² p-value' and 'X² value' columns.
- Right Panel:** History, output (a list of datasets), and a list of generated files including Pop.tab, Pop_report.tab, Pop_sub.vcf, Pop_tab_Bridge.tab, Pop_tab_DoubleDose_P1.tab, Pop_tab_DoubleDose_P2.tab, Pop_tab_SimpleDose_P1.tab, Pop_tab_SimpleDose_P2.tab, and Pop_tab_unknown.tab.

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Options passed

```
PASSED OPTIONS
--vcf: /work/GALAXY/galaxy/database/files/091/dataset_91984.dat
--MinCov: 15
--MaxCov: 300
--WinFreq: 0.01:0.1
--MinAlCov: 3
--miss: 0.05
--prefix: Pop
```

Datapoint stats

```
REPORT ON DATA POINTS
Removed line due to duplicated position in vcf file: 0
Converted to missing due to no reads: 27526 (0.9519824032316079%)
Converted to missing due to insufficient reads coverage: 14073 (0.4867125031126359%)
Converted to missing due to overcoverage: 80 (0.0027667874830534264%)
Converted to missing due to too many variant: 1621 (0.05606203137537006%)
Converted to missing due to ambiguous minority allele frequency (--WinFreq argument): 162131 (5.607275267686689%)
Converted to missing due to ambiguous minority allele coverage (--MinAlCov argument): 0 (0.0%)
```

Markers stats

```
REPORT MARKER POINT
Total marker treated: 15380
Complete line converted to missing due to unexpected vcf format (no AD or GT or DP tags): 0 (0.0%)
Marker with missing data on all individuals: 0 (0.0%)
Monomorphic makers (converted to missing for convenience): 0 (0.0%)
Di-allelic markers: 15331 (99.6814044213264%)
More than di-allelic markers (converted to missing for convenience): 49 (0.3185955786736021%)
Marker removed based on missing data cutoff (--miss argument): 4646 (30.208062418725618%)
Marker removed based on ChiSquare cutoff (--pValue argument): 1836 (11.937581274382316%)
Selected marker: 8898 (57.854356306892065%)
Marker parsed in F1 file(s): 8178
Marker parsed in unknown file(s): 163
Marker parsed in Bridge file(s): 0
Marker parsed in P2 file(s): 557
```

Statistics on the marker selection

1 Pop.tab

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Galaxy interface showing the vcfHunter tool output. A warning message at the top states: "This dataset is large and only the first megabyte is shown below." The main content is a VCF file snippet with columns for #CHROM, POS, ID, REF, ALT, QUAL, FILTER, INFO, and various FORMAT fields (P191 to P037). A red callout box highlights a specific line in the VCF file: "Vcf containing variant line passing missing data filter (not used after in this tutorial but can be of researcher interest)". On the right-hand side, the History panel shows a list of output files, with "Pop_report.tab" circled in red and labeled with a "1".

Vcf containing variant line passing missing data filter (not used after in this tutorial but can be of researcher interest)

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

The screenshot shows the Galaxy web interface with the vcfHunter tool outputs. The tool outputs are listed on the right side of the interface, including Pop.tab, Pop_report.tab, Pop_sub.vcf, Pop_tab_Bridge.tab, Pop_tab_DoubleDose_P1.tab, Pop_tab_DoubleDose_P2.tab, Pop_tab_SimpleDose_P1.tab, Pop_tab_SimpleDose_P2.tab, and Pop_tab_unknown.tab. A red circle with the number '1' and an arrow points to the Pop_sub.vcf file. A red callout box in the center of the screen contains the text: "This is the file which should contain bridge markers is empty because there are no bridge markers".

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Galaxy interface showing the output of the vcfHunter tool. The main panel displays a table of marker segregation data. A red callout box highlights a specific row with the text: "This file contained double dose markers which are heterozygous in parent P1".

Annotations on the table:

- Phase information (1 = marker has been rephased):** Points to the "rephased" column.
- Expected proportion:** Points to the "ratio" column.
- Segregation type:** Points to the "coding" column.
- Marker name (Chr+"M"+Pos):** Points to the "Marker" column.

The table headers are: Marker coding ratio rephased P191 P134 P135 P136 P137 P039 P038 P132 P133 P035 P034 P037 P036 P031 P030 P03. The rows show various marker names and their segregation patterns across different parents.

On the right side, the "output" list shows several files, with "Pop_tab_Bridge.ta" highlighted and a red circle around the number "1" in its name.

Marker coding	ratio	rephased	P191	P134	P135	P136	P137	P039	P038	P132	P133	P035	P034	P037	P036	P031	P030	P03
chr01M37131	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	--	k-	k-	hh	hh	hh	hh	hh	hh	k-	k-
chr01M37155	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	--	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M50330	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M59428	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M144740	hh,k	0.1667,0.8333	0	k-	k-	k-	hh	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M182847	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M184547	hh,k	0.1667,0.8333	0	k-	k-	k-	hh	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M184553	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M246911	hh,k	0.1667,0.8333	0	k-	k-	k-	hh	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M502178	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M664756	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M664790	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M664806	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M674819	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M681010	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M681020	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M681026	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M708929	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M768075	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M768195	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M797138	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M811010	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	hh	hh	hh	hh	hh	k-
chr01M811013	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M821045	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M823679	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M870753	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M870772	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M894164	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M894199	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M901388	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M904633	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M941445	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M951781	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M1048549	hh,k	0.1667,0.8333	0	--	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M1097161	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M1097229	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M1097277	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M1101514	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
c	h	k-	k-	k-	k-	k-	k-	k-	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh
chr01M1433061	hh,k	0.1667,0.8333	0	k-	k-	k-	k-	--	k-	k-	hh	hh	hh	hh	hh	hh	hh	k-
chr01M1461305	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	k-
chr01M1461313	hh,k	0.1667,0.8333	0	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	k-

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

This file contained double dose markers which are heterozygous in parent P2.

They are very few and this is good because P2 is diploid and thus double dose markers should not exist!

Marker	coding	ratio	rephased	P191	P134	P135	P136	P137	P039	P038	P132	P133	P035	P034	P037	P036	P031
chr01M4599854	hh, k-	0.1667, 0.8333	0	k-	k-	k-	k-	hh	k-	k-	hh	k-	k-	k-	k-	k-	k-
chr02M24813339	hh, k-	0.1667, 0.8333	0	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-	k-	k-	k-	k-
chr03M27342905	hh, k-	0.1667, 0.8333	0	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-
chr03M29789501	hh, k-	0.1667, 0.8333	0	hh	k-	k-	k-	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-
chr03M31059415	hh, k-	0.1667, 0.8333	0	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Galaxy interface showing the vcfHunter tool output. A red callout box highlights a specific file in the output list: **Pop_tab DoubleDose_P1.tab**. The callout text reads: "This file contained simple dose markers which are heterozygous in parent P1".

Marker	coding	ratio	rephased	P191	P134	P135	P136	P137	P039	P038	P132	P133	P035	P034	P037	P036	P031	P030
chr01M17228	nn,np	0.5,0.5	0	nn	np	np	np	np	np	np	--	np	np	np	np	np	np	np
chr01M20912	nn,np	0.5,0.5	0	np	nn	nn	nn	np	np	np	np	np	np	np	np	np	np	np
chr01M20916	nn,np	0.5,0.5	0	nn	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M29287	nn,np	0.5,0.5	0	np	--	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M37116	nn,np	0.5,0.5	0	nn	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M50285	nn,np	0.5,0.5	0	nn	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M50313	nn,np	0.5,0.5	0	nn	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M53409	nn,np	0.5,0.5	0	nn	np	np	np	np	np	--	np	np	np	--	np	np	np	np
chr01M53451	nn,np	0.5,0.5	0	np	nn	np	np	np	np	np	np	--	np	--	np	np	np	np
chr01M59316	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M67527	nn,np	0.5,0.5	0	nn	np	np	np	--	np	np	np	np	np	np	np	np	np	np
chr01M89917	nn,np	0.5,0.5	0	nn	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M89923	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M105158	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M127702	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M152926	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M157399	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M165977	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M165990	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M186534	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M192113	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M192120	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M192136	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M226490	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M235828	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	--	np	np	np	np
chr01M251887	nn,np	0.5,0.5	0	np	np	np	np	np	np	--	np	np	np	np	np	np	np	np
chr01M251913	nn,np	0.5,0.5	0	np	np	np	np	np	np	--	np	np	np	np	np	np	np	np
chr01M251923	nn,np	0.5,0.5	0	np	np	np	--	np	np	np	np	np	np	np	np	np	np	np
chr01M310320	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M310368	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M326767	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M340656	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M381756	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M389536	nn,np	0.5,0.5	0	np	np	np	np	--	np	np	np	np	np	np	np	np	np	np
chr01M408718	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M415461	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M466101	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M471076	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M471092	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M488027	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M502137	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M502140	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M502151	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M511685	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	--	np	np	np
chr01M512201	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M528716	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M534858	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M548511	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M563312	nn,np	0.5,0.5	0	np	np	np	np	np	np	np	np	np	np	np	np	np	np	np

Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Marker coding	ratio	rephased	P191	P134	P135	P136	P137	P039	P038	P132	P133	P035	P034	P037	P036
chr01M2909259	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3084769	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3230825	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3284968	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3285013	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3358254	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3395458	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3417999	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3493115	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3507847	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3580672	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3758744	nn, np	0.5, 0.5 0	np	np	np	--	np	np	np	np	np	np	np	np	np
chr01M3775407	np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	--	np	np	np
chr01M3775423	nn, np	0.5, 0.5 0	--	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3837583	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3837669	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3847042	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3856981	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3881851	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3901090	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3901141	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3913984	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3926703	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M3958080	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4015771	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4032506	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4047381	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4047390	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4053672	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4127313	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4145915	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4150482	nn, np		np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4163170	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4167085	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4185588	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4222080	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4235526	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4242788	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	--	np	np
chr01M4242794	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	--	np	np
chr01M4269479	nn, np	0.5, 0.5 0	--	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4288400	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	--	np	np
chr01M4288430	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4325073	nn, np	0.5, 0.5 0	np	np	np	np	np	--	np	np	np	np	--	np	np
chr01M4334576	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4366480	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4426286	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4426367	nn, np	0.5, 0.5 0	--	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4426401	nn, np	0.5, 0.5 0	--	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4435206	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4435217	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4435339	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4526241	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4537287	nn, np	0.5, 0.5 0	np	np	np	--	np	np	np	np	np	np	np	np	np
chr01M4557883	nn, np	0.5, 0.5 0	--	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4577078	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	--	np	np
chr01M4598287	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4616823	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4630312	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4668683	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M4692503	nn, np	0.5, 0.5 0	np	np	np	np	np	np	np	np	np	np	--	np	np

This file contained simple dose markers which are heterozygous in parent P2



Genetic mapping analysis with vcfHunter tool

Selecting segregating markers in vcf files: outputs

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Admin | Help | User | Using 19%

Tools | **vcfHunter** | VCF Filter | vcf2allPropAndCov | vcf2allPropAndCovByChr | vcf2popNew | RecombCalculatorDDose | Draw_dot_plot | KDE_classifier | METAGENOMICS | FROGS | EVOLUTION/PHYLOGENY | Comparative Genomics | NCBI BLAST+ | Genfam | Protein analyses | STATISTICS/GRAPHICS | Statistics | Graph/Display Data | SOUTHGREEN PROJECTS | SNPPlay3 | GNPAnnot Tools | GNPAnnot Converters | ESTtik | Expression data | SAT

Marker coding	ratio	rephased	P191	P134	P135	P136	P137	P039	P038	P132	P133	P035	P034	P037	P036
chr01M1455285	nn,np	0.5,0.5 0	nn	np	np	np	np	np	nn	nn	np	nn	np	np	np
chr01M1455292	nn,np	0.5,0.5 0	nn	np	np	np	np	np	nn	nn	np	nn	np	np	np
chr01M1455306	nn,np	0.5,0.5 0	nn	np	np	np	np	np	nn	nn	np	nn	np	np	np
chr01M3288418	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M3592725	hh,k-	0.1667,0.8333	0	--	k-	k-	k-	k-	k-	hh	k-	--	k-	k-	k-
chr01M3603814	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	hh	k-	hh	k-
chr01M3603846	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	hh	k-	hh	k-
chr01M3754938	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-	--	k-
chr01M4325104	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	--	hh	k-	k-	k-
chr01M4372486	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-
chr01M4491444	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	hh	k-	k-
chr01M44537232	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M44537284	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M4630303	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M4668632	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	hh	k-	k-
chr01M4674927	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	hh	k-	k-
chr01M4848289	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	hh	k-	k-
chr01M4938144	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M4996277	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5037471	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5280383	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5376348	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5553519	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5639249	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5846424	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5874383	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5893880	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5893907	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M5913290	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M6148149	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M6283060	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M6307631	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M6945825	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M6945889	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M7536345	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M7536847	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M7793878	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M7849181	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	hh	k-	hh	k-	k-
chr01M7970956	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	--	hh	k-	hh	k-	k-
chr01M7970982	nn,np	0.5,0.5 0	np	np	np	np	np	np	np	np	np	np	np	np	np
chr01M8847701	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-	hh	hh
chr01M9309560	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	k-	hh	k-	hh	hh
chr01M9630679	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-
chr01M9885247	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M9885251	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M991162	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M10373361	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M10373388	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	hh	k-	hh	k-	k-
chr01M10373414	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	hh	k-	hh	k-	k-
chr01M10393063	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	hh	k-	hh	k-	k-
chr01M10521951	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M11010230	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	hh	k-	--	k-	k-
chr01M11446058	nn,np	0.5,0.5 0	--	np	np	np	np	np	np	np	np	np	np	--	np
chr01M11192384	nn,np	0.5,0.5 0	nn	np	np	np	np	np	np	np	np	np	--	np	np
chr01M11747728	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-
chr01M12845937	hh,k-	0.1667,0.8333	0	hh	k-	k-	k-	k-	k-	hh	hh	k-	k-	hh	hh
chr01M13011194	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-
chr01M13011203	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-	k-
chr01M13125723	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	--	k-	k-	k-	k-	k-
chr01M14434809	hh,k-	0.1667,0.8333	0	k-	k-	k-	k-	k-	k-	hh	k-	k-	k-	k-	k-

History | < Back to VCFHunter | output | a list of datasets | Pop.tab | Pop_report.tab | Pop_sub.vcf | Pop_tab_Bridge.tab | Pop_tab_DoubleDose_P1.tab | Pop_tab_DoubleDose_P2.tab | Pop_tab_SimpleDose_P1.tab | **Pop_tab_SimpleDose_P12.tab** | Pop_tab_unknown.tab | View data

This file contained markers for which the heterozygous status of parent P1 and P2 is unclear and thus they could not be attributed to one of the parents

Genetic mapping analysis with vcfHunter tool

Calculating marker segregation distortion on parent P2 simple dose markers

The screenshot shows the Galaxy web interface. On the left, a sidebar lists various tools, with 'RecombCalculatorDDose' highlighted and a red circle containing the number '1' next to it. The main panel displays the configuration for the 'RecombCalculatorDDose' tool (Galaxy Version 0.1.0). The configuration includes:

- The marker file matrix:** 125: Kunnan_stats.tab
- Are marker phased?:** n
- Analysis to perform:** R (Calculate recombination rate / S: Calculate segregation distortions)

Below the configuration, there is an 'Execute' button and a list of authors and support information. The tool's description and inputs/outputs are also visible.

RecombCalculatorDDose

Description

This program perform is designed to calculate frequencies of recombination observed between two pairs of markers. It can also calculate marker segregation distortion.

Inputs:

The marker file matrix = Pop_tab_SegregationName_Parent.tab from vcf2popNew

Output:

REC.tab: A tabulated file of pairwise marker recombination rate (analysis to perform R).
SegDist.tab: A tabulated file of marker segragtion distortions (analysis to perform S).

Citations Show BibTeX

Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

History

< Back to VCFHunter

output

a list of datasets

- Pop.tab
- Pop_report.tab
- Pop_sub.vcf
- Pop_tab_Bridge.tab
- Pop_tab_DoubleDose_P1.tab
- Pop_tab_DoubleDose_P2.tab
- Pop_tab_SimpleDose_P1.tab
- Pop_tab_SimpleDose_P2.tab
- Pop_tab_unknown.tab

Genetic mapping analysis with vcfHunter tool

Calculating marker segregation distortion on parent P2 simple dose markers

RecombCalculatorDDose (Galaxy Version 0.1.0)

The marker file matrix

125: Kunnan_stats.tab

output of vcf2pop

Are marker phase

125: Kunnan_stats.tab

124: Kunnan_AlleleOriginAndRatio.tab

123: Kunnan_Ratio.png

Analysis to perform

R

122: Kunnan_Cov.png

120: Origin.tab

R: Calculate recombination

119: chr09_test.vcf

118: chr04_test.vcf

117: chr03_test.vcf

Execute

RecombCalculatorDDose

Description

This program perform is designed to calculate frequencies of recombination

Inputs:

The marker file matrix = Pop_tab_SegregationName_Parent.tab from

Output:

REC.tab: A tabulated file of pairwise marker recombination rate (a

SegDist.tab: A tabulated file of marker segregation distortions (ana

Citations Show BibTeX

Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination in *Evolution*, [doi:10.1093/molbev/msy199][Link]

File which are in collections are not available through by galaxy workflow!!!

The problem is that we want to access the "Pop_tab_SimpleDose_P2.tab" file...

To solve this problem, we have to:

- 1- get this file on our computer
- 2- upload this file on galaxy but this time not as part of a collection

Genetic mapping analysis with vcfHunter tool

1- Downloading the Pop_tab_SimpleDose_P2.tab

The screenshot shows the Galaxy web interface with the **RecombCalculatorDDose** tool configuration page. The tool is set to use the marker file matrix `125: Kunnan_stats.tab` and is configured to perform analysis `R` (Calculate recombination rate). A dialog box titled "Ouverture de Galaxy145-[Pop_tab_SimpleDose_P2.tab].null" is open, showing options to open the file. The "Enregistrer le fichier" option is selected, and the "OK" button is highlighted. The right sidebar shows a list of datasets, with `Pop_tab_SimpleDose_P2.tab` highlighted and its download icon circled in red. The main content area includes the tool's description, inputs, and output details.

RecombCalculatorDDose (Galaxy Version 0.1.0)

The marker file matrix
125: Kunnan_stats.tab
output of vcf2popNew

Are marker phased?
n

Analysis to perform
R
R: Calculate recombination rate / S: Calculate segregation distortions

Execute

Author Guillaume MARTIN (quillaume.martin@cirad.fr)
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

Description
This program perform is designed to calculate frequencies of recombination observed between two pairs of markers. It can also calculate marker segregation distortion.

Inputs:
The marker file matrix = Pop_tab_SegregationName_Parent.tab from vcf2popNew

Output:
REC.tab: A tabulated file of pairwise marker recombination rate (analysis to perform R).
SegDist.tab: A tabulated file of marker segragtion distortions (analysis to perform S).

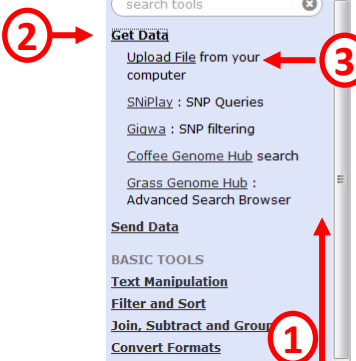
Citations [Show BibTeX](#)
Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

History
Back to VCFHunter
output
a list of datasets

- Pop.tab
- Pop_report.tab
- Pop_sub.vcf
- Pop_tab_Bridqe.tab
- Pop_tab_DoubleDose_P1.tab
- Pop_tab_DoubleDose_P2.tab
- Pop_tab_SimpleDose_P1.tab
- Pop_tab_SimpleDose_P2.tab** (database: ?)
- Pop_tab_unknown.tab

Genetic mapping analysis with vcfHunter tool

1- Uploading the Pop_tab_SimpleDose_P2.tab onto Galaxy



Galaxy interface showing the RecombCalculatorDDose tool configuration. The tool is set to use the marker file matrix '125: Kunnan_stats.tab' and is configured to calculate recombination rate (R). The output list on the right includes 'Pop.tab', 'Pop_report.tab', 'Pop_sub.vcf', 'Pop_tab_Bridge.tab', 'Pop_tab_DoubleDose_P1.tab', 'Pop_tab_DoubleDose_P2.tab', 'Pop_tab_SimpleDose_P1.tab', 'Pop_tab_SimpleDose_P2.tab', and 'Pop_tab_unknown.tab'.

RecombCalculatorDDose (Galaxy Version 0.1.0)

The marker file matrix
125: Kunnan_stats.tab
output of vcf2popNew

Are marker phased?
n

Analysis to perform
R
R: Calculate recombination rate / S: Calculate segregation distortions

Execute

Author Guillaume MARTIN (quillaume.martin@cirad.fr)
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

RecombCalculatorDDose

Description
This program perform is designed to calculate frequencies of recombination observed between two pairs of markers. It can also calculate marker segregation distortion.

Inputs:
The marker file matrix = Pop_tab_SegregationName_Parent.tab from vcf2popNew

Output:
REC.tab: A tabulated file of pairwise marker recombination rate (analysis to perform R).
SegDist.tab: A tabulated file of marker segregation distortions (analysis to perform S).

Citations Show BibTeX
Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

Genetic mapping analysis with vcfHunter tool

1- Uploading the Pop_tab_SimpleDose_P2.tab onto Galaxy

Galaxy145-[Pop_tab_SimpleDose_P2.tab].null

Name	Size	Type	Genome	Settings	Status
Galaxy145-[Pop_tab_SimpleDose_P2.tab].null	321.5 KB	Auto-det...	unspecified (?)		100%

Start

1
Drag and drop the downloaded file

2
3

Genetic mapping analysis with vcfHunter tool

Calculating marker segregation distortion on parent P2 simple dose markers

The screenshot shows the Galaxy web interface with the **RecombCalculatorDDose** tool (Galaxy Version 0.1.0) configured. The tool options are as follows:

- The marker file matrix:** 147: Galaxy145-[Pop_tab_SimpleDose_P2.tab].null (circled 1)
- Are marker phased?:** n (circled 2)
- Analysis to perform:** S (circled 3)
- Execute button:** (circled 4)

Annotations on the right side of the tool interface:

- A red dashed arrow points from the circled '1' to the file name in the History panel.
- A red arrow points from the circled '2' to the letter 'n'.
- A red arrow points from the circled '3' to the letter 'S'.

The History panel on the right shows the following datasets:

- VCFHunter (16 shown, 112 deleted, 19 hidden)
- 147: Galaxy145-[Pop_tab_SimpleDose_P2.tab].null (highlighted with a red arrow)
- 137: output
- Kunnan AlleleOriginAndRatio.tab
- 123: Kunnan_Ratio.png
- 122: Kunnan_Cov.png
- 121: vcf.conf (a list of 5 datasets)
- 120: Origin.tab
- 119: chr09_test.vcf
- 118: chr04_test.vcf
- 117: chr03_test.vcf
- 116: chr02_test.vcf
- 115: chr01_test.vcf
- 114: Carto.vcf
- 113: CartoRef.app

A white box with a black border contains the text: **Downloaded file appears**, with a red arrow pointing to the highlighted dataset in the History panel.

RecombCalculatorDDose

Description

This program perform is designed to calculate frequencies of recombination observed between two pairs of markers. It can also calculate marker segregation distortion.

Inputs:

The marker file matrix = Pop_tab_SegregationName_Parent.tab from vcf2popNew

Output:

REC.tab: A tabulated file of pairwise marker recombination rate (analysis to perform R).
SegDist.tab: A tabulated file of marker segragtion distortions (analysis to perform S).

Citations [Show BibTeX](#)

Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

Genetic mapping analysis with vcfHunter tool

Calculating marker segregation distortion : output

A two column file with:

- 1- marker name
- 2- marker segregation distortion

1	2
chr01M2909259	1.0675463526216225
chr01M3084769	1.0587658909783821
chr01M3230825	1.123551035124867
chr01M3284968	0.4798874785587451
chr01M3285013	0.5276568002953714
chr01M3358254	0.8080719000803791
chr01M3395458	1.1328964223140199
chr01M3417999	1.0675463526216225
chr01M3493115	0.9336553965234122
chr01M3507847	0.5276568002953714
chr01M3580672	0.5726718245857745
chr01M3758744	0.9336553965234122
chr01M3775407	1.0675463526216225
chr01M3775423	0.34098131172920304
chr01M3837583	0.5276568002953714
chr01M3837669	0.43051680529937114
chr01M3847042	1.0587658909783821
chr01M3856981	0.4763798150217415
chr01M3881851	0.523767198186683
chr01M3901090	0.523767198186683
chr01M3901141	0.4763798150217415
chr01M3913984	0.8594678641097688
chr01M3926703	0.995462582229307
chr01M3958080	0.4763798150217415
chr01M4015771	0.8080719000803791
chr01M4032506	1.0587658909783821
chr01M4047381	1.1424366088102702
chr01M4047390	1.1424366088102702
chr01M4053672	1.1328964223140199
chr01M4127313	0.4798874785587451
chr01M4145915	0.5276568002953714
chr01M4150482	0.523767198186683
chr01M4163170	0.6277828140473716
chr01M4167085	1.1328964223140199
chr01M4185588	1.123551035124867
chr01M4222080	0.63257328864087
chr01M4235526	0.7454217266286189
chr01M4242788	0.995462582229307
chr01M4242794	1.0675463526216225
chr01M4269479	0.5276568002953714
chr01M4288400	0.5276568002953714
chr01M4288430	1.050163073656123
chr01M4325073	0.4798874785587451

Genetic mapping analysis with vcfHunter tool

Calculating marker segregation distortion : output

A two column file with:

- 1- marker name
- 2- marker segregation distortion

1	2
chr01M2909259	1.0675463526216225
chr01M3084769	1.0587658909783821
chr01M3230825	1.123551035124867
chr01M3284968	0.4798874785587451
chr01M3285013	0.5276568002953714
chr01M3358254	0.8080719000803791
chr01M3395458	1.1328964223140199
chr01M3417999	1.0675463526216225
chr01M3493115	0.9336553965234122
chr01M3507847	0.5276568002953714
chr01M3580672	0.5726718245857745
chr01M3758744	0.9336553965234122
chr01M3775407	1.0675463526216225
chr01M3775423	0.34098131172920304
chr01M3837583	0.5276568002953714
chr01M3837669	0.43051680529937114
chr01M3847042	1.0587658909783821
chr01M3856981	0.4763798150217415
chr01M3881851	0.523767198186683
chr01M3901090	0.523767198186683
chr01M3901141	0.4763798150217415
chr01M3913984	0.8594678641097688
chr01M3926703	0.995462582229307
chr01M3958080	0.4763798150217415
chr01M4015771	0.8080719000803791
chr01M4032506	1.0587658909783821
chr01M4047381	1.1424366088102702
chr01M4047390	1.1424366088102702
chr01M4053672	1.1328964223140199
chr01M4127313	0.4798874785587451
chr01M4145915	0.5276568002953714
chr01M4150482	0.523767198186683
chr01M4163170	0.6277828140473716
chr01M4167085	1.1328964223140199
chr01M4185588	1.123551035124867
chr01M4222080	0.63257328864087
chr01M4235526	0.7454217266286189
chr01M4242788	0.995462582229307
chr01M4242794	1.0675463526216225
chr01M4269479	0.5276568002953714
chr01M4288400	0.5276568002953714
chr01M4288430	1.050163073656123
chr01M4325073	0.4798874785587451

Genetic mapping analysis with vcfHunter tool

Calculating pairwise marker recombination rate on P2 simple dose

The screenshot shows the Galaxy web interface. On the left, a sidebar lists various tools, with 'RecombCalculatorDDose' highlighted and a red circle containing the number '1' next to it. The main panel displays the configuration for the 'RecombCalculatorDDose' tool (Galaxy Version 0.1.0). The configuration includes:

- The marker file matrix:** 125: Kunnan_stats.tab
- Are marker phased?:** n
- Analysis to perform:** R (Calculate recombination rate / S: Calculate segregation distortions)

Below the configuration, there is an 'Execute' button and a list of authors and support information. The tool's description and inputs/outputs are also visible.

RecombCalculatorDDose

Description

This program perform is designed to calculate frequencies of recombination observed between two pairs of markers. It can also calculate marker segregation distortion.

Inputs:

The marker file matrix = Pop_tab_SegregationName_Parent.tab from vcf2popNew

Output:

REC.tab: A tabulated file of pairwise marker recombination rate (analysis to perform R).
SegDist.tab: A tabulated file of marker segragtion distortions (analysis to perform S).

Citations [Show BibTeX](#)

Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

History

< Back to VCFHunter

output

a list of datasets

- Pop.tab
- Pop_report.tab
- Pop_sub.vcf
- Pop_tab_Bridge.tab
- Pop_tab_DoubleDose_P1.tab
- Pop_tab_DoubleDose_P2.tab
- Pop_tab_SimpleDose_P1.tab
- Pop_tab_SimpleDose_P2.tab
- Pop_tab_unknown.tab

Genetic mapping analysis with vcfHunter tool

Calculating pairwise marker recombination rate on P2 simple dose

The screenshot shows the Galaxy web interface with the **RecombCalculatorDDose** tool (Galaxy Version 0.1.0) configured. The tool options are as follows:

- The marker file matrix:** 147: Galaxy145-[Pop_tab_SimpleDose_P2.tab].null (circled 1)
- Are marker phased?:** n (circled 2)
- Analysis to perform:** R (circled 3)
- Execute button:** (circled 4)

Red arrows point from the circled '1' to the 'VCFHunter' dataset in the History panel and from the circled '2' to the 'n' parameter. The History panel on the right lists various datasets, including 'VCFHunter' (160.58 MB) and several 'test.vcf' files.

RecombCalculatorDDose Description:
This program perform is designed to calculate frequencies of recombination observed between two pairs of markers. It can also calculate marker segregation distortion.

Inputs:
The marker file matrix = Pop_tab_SegregationName_Parent.tab from vcf2popNew

Output:
REC.tab: A tabulated file of pairwise marker recombination rate (analysis to perform R).
SegDist.tab: A tabulated file of marker segragtion distortions (analysis to perform S).

Citations: Show BibTeX
Baurens, Franc-Christophe and Martin, Guillaume and Hervouet, Catherine and Salmon, Frédéric and Yohomé, David and Ricci, Sébastien and Rouard, Mathieu and Habas, Remy and Lemainque, Arnaud and Yahiaoui, Nabila and et al. (2018). Recombination and large structural variations shape interspecific edible bananas genomes. In *Molecular Biology and Evolution*, [doi:10.1093/molbev/msy199][Link]

Genetic mapping analysis with vcfHunter tool

Calculating pairwise marker recombination rate: output

A square matrix of marker pairwise recombination rate

ID	chr01M2909259	chr01M3084769	chr01M320825	chr01M3284968	chr01M3285013	chr01M3358254	chr01M3395458	chr01M3417999	chr01M3493115	chr01M35078
chr01M2909259	0.0	0.0109090909090909	0.021897810218978103	0.046511627906976744	0.028735632183908046	0.046242774566473986	0.05020312138728324	0.02583025830258302	0.01454545454545454	0.00359712230215827
chr01M3084769	0.0109090909090909	0.0	0.01079136690647482	0.011428571428571429	0.0	0.0	0.03482758620689655	0.01454545454545454	0.00359712230215827	
chr01M320825	0.021897810218978103	0.01079136690647482	0.0	0.011428571428571429	0.0	0.0	0.005780346820809248	0.005813953488372093	0.0	
chr01M3284968	0.046511627906976744	0.028735632183908046	0.011428571428571429	0.0	0.0	0.0	0.005780346820809248	0.005813953488372093	0.005813953488372093	
chr01M3285013	0.046242774566473986	0.028735632183908046	0.011428571428571429	0.0	0.0	0.0	0.005780346820809248	0.005813953488372093	0.005813953488372093	
chr01M3358254	0.0520312138728324	0.03482758620689655	0.017142857142857144	0.0	0.0	0.0	0.01156069364168497	0.011627906976744186	0.011627906976744186	
chr01M3395458	0.025830258302583026	0.025830258302583026	0.014545454545454545	0.0035971223021582736	0.005780346820809248	0.005780346820809248	0.005780346820809248	0.005780346820809248	0.005780346820809248	
chr01M3417999	0.026022304832713755	0.014598540145985401	0.0036231884057971015	0.005813953488372093	0.005780346820809248	0.005780346820809248	0.005780346820809248	0.005780346820809248	0.005780346820809248	
chr01M3493115	0.026022304832713755	0.014598540145985401	0.0036231884057971015	0.005813953488372093	0.005780346820809248	0.005780346820809248	0.005780346820809248	0.005780346820809248	0.005780346820809248	
chr01M35078	0.0520312138728324	0.03482758620689655	0.017142857142857144	0.0	0.0	0.0	0.011494252873563218	0.01156069364168497	0.01156069364168497	
chr01M350672	0.056818181818181816	0.039327272727272727	0.039327272727272727	0.0	0.0	0.0	0.00392156862745098	0.040697674418604654	0.040697674418604654	
chr01M3758744	0.0449438202247191	0.0449438202247191	0.033027272727272727	0.0	0.0	0.0	0.040697674418604654	0.040697674418604654	0.040697674418604654	
chr01M3775407	0.04477611940298507	0.04477611940298507	0.062827272727272727	0.0	0.0	0.0	0.040697674418604654	0.040697674418604654	0.040697674418604654	
chr01M3775423	0.08092485549132948	0.08092485549132948	0.062827272727272727	0.0	0.0	0.0	0.0199203187250996	0.0199203187250996	0.0199203187250996	
chr01M3837583	0.08670520231213873	0.08670520231213873	0.068527272727272727	0.0	0.0	0.0	0.024	0.04597701149425287	0.04597701149425287	
chr01M3837669	0.07514450867052024	0.07514450867052024	0.068527272727272727	0.0	0.0	0.0	0.024	0.04597701149425287	0.04597701149425287	
chr01M3847092	0.04979047970479705	0.04979047970479705	0.068527272727272727	0.0	0.0	0.0	0.04285714	0.02189781021897810	0.02189781021897810	
chr01M3856548	0.08045977011494253	0.08045977011494253	0.062527272727272727	0.0	0.0	0.0	9523808	0.04571428571428571	0.04571428571428571	
chr01M3881851	0.08571428571428572	0.08571428571428572	0.067759852727272727	0.0	0.0	0.0	9523808	0.04545454545454545	0.04545454545454545	
chr01M3901090	0.08522727272727272	0.08522727272727272	0.067759852727272727	0.0	0.0	0.0	0.02390438247011952	0.02390438247011952	0.02390438247011952	
chr01M3901141	0.08571428571428572	0.08571428571428572	0.06818181818181818	0.05084745762711865	0.023715415019762844	0.02362204724409488	0.024	0.04571428571428571	0.04571428571428571	
chr01M3913984	0.05185185185185185	0.05185185185185185	0.04	0.028985507246376812	0.045714285714285714	0.04545454545454545	0.05142857142857143	0.02554744525547445	0.02554744525547445	
chr01M3926703	0.05223880597014925	0.05223880597014925	0.040444117647058284	0.029197080291970802	0.046242774566473986	0.04597701149425287	0.05202312138728324	0.05202312138728324	0.05202312138728324	
chr01M3950580	0.08620689655172414	0.08620689655172414	0.07386363636363637	0.0564971751424294	0.02766798418972332	0.02755905511810236	0.028	0.05142857142857143	0.05142857142857143	
chr01M4015771	0.052830188679245285	0.052830188679245285	0.040892193308550186	0.02952029520295203	0.046511627906976744	0.046242774566473986	0.05232558139534884	0.05232558139534884	0.05232558139534884	
chr01M4032506	0.04814814814814815	0.04814814814814815	0.03676470588235294	0.025362318840579712	0.04597701149425287	0.045714285714285714	0.05142857142857143	0.05142857142857143	0.05142857142857143	
chr01M4047381	0.05639097744360902	0.05639097744360902	0.04460966542750929	0.03308823529411765	0.05232558139534884	0.05202312138728324	0.05232558139534884	0.05232558139534884	0.05232558139534884	
chr01M4047390	0.05639097744360902	0.05639097744360902	0.04460966542750929	0.03308823529411765	0.05232558139534884	0.05202312138728324	0.05232558139534884	0.05232558139534884	0.05232558139534884	
chr01M4053672	0.052434456928838954	0.052434456928838954	0.04059040590405904	0.02909090909090909	0.046242774566473986	0.04597701149425287	0.05202312138728324	0.05202312138728324	0.05202312138728324	
chr01M4127313	0.10465116279069768	0.10465116279069768	0.08620689655172414	0.06857142857142857	0.038585673705179286	0.038585673705179286	0.03643724696356275	0.03643724696356275	0.03643724696356275	
chr01M4145915	0.1040624277456648	0.1040624277456648	0.08571428571428572	0.06818181818181818	0.03571428571428571	0.03571428571428571	0.036290322580645164	0.036290322580645164	0.036290322580645164	
chr01M4150482	0.10857142857142857	0.10857142857142857	0.0903954802259887	0.07303370786516854	0.03543307086614173	0.0392156862745098	0.04	0.06818181818181818	0.06818181818181818	
chr01M4163170	0.10285714285714286	0.10285714285714286	0.0847457627118644	0.06741573033707865	0.03557312252964427	0.03543307086614173	0.036	0.0625	0.0628571428	
chr01M4167085	0.0599250936329588	0.0599250936329588	0.04814814814814815	0.0366303663003663	0.0574126436781609	0.05714285714285714	0.057803468208092484	0.057803468208092484	0.057803468208092484	
chr01M4185588	0.05904059040590406	0.05904059040590406	0.04744525547445255	0.036231884057971016	0.05714285714285714	0.05714285714285714	0.06285714285714286	0.06285714285714286	0.06285714285714286	
chr01M4222080	0.1040624277456648	0.1040624277456648	0.08571428571428572	0.06779661106494153	0.038585673705179286	0.038585673705179286	0.03571428571428571	0.036290322580645164	0.036290322580645164	
chr01M4235262	0.097770114942528736	0.097770114942528736	0.08	0.0625	0.038585673705179286	0.03571428571428571	0.036585365853658534	0.05747126436781609	0.057803468208092484	
chr01M4242788	0.06037735849056604	0.06037735849056604	0.04814814814814815	0.03676470588235294	0.057803468208092484	0.05747126436781609	0.0635938150290173	0.0635938150290173	0.0635938150290173	
chr01M4242794	0.06060606060606061	0.06060606060606061	0.048321375464688404	0.03690303690369004	0.05813953488372093	0.057803468208092484	0.0635938150290173	0.0635938150290173	0.0635938150290173	
chr01M4269479	0.1040624277456648	0.1040624277456648	0.08571428571428572	0.06818181818181818	0.038585673705179286	0.038585673705179286	0.03571428571428571	0.036290322580645164	0.036290322580645164	
chr01M4288400	0.0982689595375723	0.0982689595375723	0.08	0.0625	0.038585673705179286	0.03571428571428571	0.03585673705179286	0.036290322580645164	0.036290322580645164	
chr01M4288430	0.05904059040590406	0.05904059040590406	0.04727272727272727	0.036101083032490974	0.056818181818181816	0.05649717514124294	0.0625	0.0326086956521391	0.0326086956521391	
chr01M4325073	0.10465116279069768	0.10465116279069768	0.08620689655172414	0.06857142857142857	0.038585673705179286	0.03571428571428571	0.03643724696356275	0.03643724696356275	0.03643724696356275	
chr01M4334576	0.0559701149253713145	0.0559701149253713145	0.044117647058823529	0.03296703296703297	0.05232558139534884	0.05202312138728324	0.057803468208092484	0.057803468208092484	0.057803468208092484	
chr01M4366480	0.09883720930232558	0.09883720930232558	0.08620689655172414	0.06857142857142857	0.038585673705179286	0.038585673705179286	0.036290322580645164	0.036290322580645164	0.036290322580645164	
chr01M4426286	0.09883720930232558	0.09883720930232558	0.08045977011494253	0.06285714285714286	0.036	0.038585673705179286	0.036734693877551024	0.057803468208092484	0.057803468208092484	
chr01M4426367	0.09714285714285714	0.09714285714285714	0.0847457627118644	0.06741573033707865	0.03529411764705882	0.03529411764705882	0.036	0.0625	0.0628571428	
chr01M4426401	0.09714285714285714	0.09714285714285714	0.0847457627118644	0.06741573033707865	0.03529411764705882	0.03529411764705882	0.036	0.0625	0.0628571428	
chr01M4435206	0.08525825825825826	0.08525825825825826	0.04323232323232327	0.03523232323232327	0.05818181818181816	0.05648717514124294	0.0635	0.03232323232323233	0.03232323232323233	

Genetic mapping analysis with vcfHunter tool

Plotting pairwise recombination

Galaxy Version 0.1.0

The pairwise matrix marker file
150: REC.tab
generated by RecombCalculatorDDose (REC.tab)

I already have a locus file
yes

Loci to plot with their locations
150: REC.tab

List of chromosomes to draw in this order
separated by ':'. If not filled, all chromosomes are used.

Agg file locating scaffolds in the reference sequence
Nothing selected

stats
Nothing selected

A value specifying if the marker position should be defined based on physical position or not
n

Execute

Author Guillaume MARTIN (guillaume.martin@cirad.fr)
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

Draw_dot_plot
Description
This program draw a dotplot based on marker pairwise recombination file obtained from RecombCalculatorDDose.

Inpputs:
The pairwise matrix marker file = REC.tab from RecombCalculatorDDose
Stats = SegDist.tab from RecombCalculatorDDose

Outputs:
A dotplot file representing pairwise marker linkage

History
search datasets
VCFHunter
18 shown, 113 deleted, 19 hidden
166.13 MB
150: REC.tab
148: SegDist.tab
137: output
126: output
125: Kunnan_stats.tab
124: Kunnan_AlleleOriginAndRatio.tab
123: Kunnan_Ratio.png
122: Kunnan_Cov.png
121: vcf.conf
120: Origin.tab
119: chr09_test.vcf
118: chr04_test.vcf
117: chr03_test.vcf
116: chr02_test.vcf
115: chr01_test.vcf
114: Carto.vcf
113: CartoRef.agg

Genetic mapping analysis with vcfHunter tool

Plotting pairwise recombination: outputs

The screenshot displays the Galaxy web interface. The main area shows a heatmap visualization of pairwise recombination, with a color scale from red (high recombination) to blue (low recombination) across a genomic region from 0.0 to 0.5. A red callout box with the number '1' points to the 'output_heatmap.png' file in the History panel on the right. A red callout box with the text 'Color code of the heatmap' is positioned below the heatmap.

Genetic mapping analysis with vcfHunter tool

Plotting pairwise recombination: outputs

• Marker linkage
• Marker ordered with no physical distances (i.e. same distance between each point)

Galaxy interface showing a pairwise recombination heatmap. The heatmap is a lower triangular matrix with chromosomes labeled on the axes: chr01, chr02, and chr03. The color scale ranges from blue (low recombination) to red (high recombination). A red callout box highlights the heatmap with the text: "• Marker linkage" and "• Marker ordered with no physical distances (i.e. same distance between each point)".

History panel (right side) showing a list of datasets. A red circle with the number 1 highlights the dataset "152: output_heatmap.png". Other datasets include "151: output.png", "150: REC.tab", "148: SeqDist.tab", "147: Galaxy145-[Pop tab SimpleDose P2.tab].null", "137: output", "126: output", "125: Kunnan_stats.tab", "124: Kunnan AlleleOriginAndRatio.tab", "123: Kunnan_Ratio.png", "122: Kunnan_Cov.png", "121: vcf.conf", "120: Origin.tab", "119: chr09_test.vcf", "118: chr04_test.vcf", "117: chr03_test.vcf", "116: chr02_test.vcf", and "115: chr01_test.vcf".

Genetic mapping analysis with vcfHunter tool

Plotting pairwise recombination with other options

The screenshot displays the Galaxy web interface. The main workspace shows a pairwise recombination heatmap with a color scale from blue (low recombination) to red (high recombination). The heatmap is divided into three sections labeled chr01, chr02, and chr03. The left sidebar contains a list of tools, with 'vcfHunter' selected. The right sidebar shows the 'History' panel, listing various datasets. A red circle with the number '1' highlights the 'vcfHunter' entry in the history panel.

Tools List:

- Rice Variant Analysis (Rice 3k, IRIGIN, High Density Rice Array (HDRA, 700k SNPs))
- GENOME HARVEST
- TransPo-RG Transfer of Position to Resequenced Genome
- parental SNP - Detect parental SNP of hybrids
- Visualization
- TraceAncestor
- vcfHunter
 - VCF Filter
 - vcf2allPropAndCov
 - vcf2allPropAndCovByChr
 - vcf2popNew
 - RecombCalculatorDDose
 - Draw_dot_plot
- KDE_classifier
- METAGENOMICS
- FROGS
- EVOLUTION/PHYLOGENY
- Comparative Genomics
- NCBI BLAST+
- Genfam
- Protein analyses
- STATISTICS/GRAPHICS
- Statistics
- Graph/Display Data
- SOUTHGREEN PROJECTS
- SNIPlay3
- GNPAnnot Tools
- GNPAnnot Converters
- ESThik
- Expression data
- SAT
- CMAEE tools

History Panel:

- search datasets
- VCfHunter
 - 20 shown, 113 deleted, 19 hidden
 - 166.26 MB
 - 0: histmap.png
 - 151: output.png
 - 150: REC.tab
 - 148: SeqDist.tab
 - 147: Galaxy145-[Pop_tab SimpleDose P2.tab], null
 - 137: output
 - 126: output
 - 125: Kunnan_stats.tab
 - 124: Kunnan_AlleleOriginAndRatio.tab
 - 123: Kunnan_Ratio.png
 - 122: Kunnan_Cov.png
 - 121: vcf.conf
 - 120: Origin.tab
 - 119: chr09_test.vcf
 - 118: chr04_test.vcf
 - 117: chr03_test.vcf
 - 116: chr02_test.vcf
 - 115: chr01_test.vcf

Genetic mapping analysis with vcfHunter tool

Plotting pairwise recombination with other options

Draw_dot_plot (Galaxy Version 0.1.0)

The pairwise matrix marker file
150: REC.tab
generated by RecombCalculatorDDose (REC.tab)

I already have a locus file
no
col1: marker name, col2: chromosome, col3: position

Pop_tab_SegregationName_Parent.tab from vcf2popNew
147: Galaxy145-[Pop_tab_SimpleDose_P2.tab].null
create your locus file

List of chromosomes to draw in this order
separated by ':'. If not filled, all chromosomes are used.

Agg file locating scaffolds in the reference sequence
113: CartoRef.agp

stats
148: SegDist.tab
A two column file with column 1: marker name, column2: statistics (SegDist.tab)

A value specifying if the marker position should be defined based on physical position or not
y

Execute

Author Guillaume MARTIN (guillaume.martin@cirad.fr)
Galaxy integration Aurore Comte
Support For any questions about Galaxy integration, please send an e-mail to aurore.comte@ird.fr

Draw_dot_plot
Description
This program draw a dotplot based on marker pairwise recombination file obtained from RecombCalculatorDDose.

Inputs:
The pairwise matrix marker file = REC.tab from RecombCalculatorDDose
Stats = SegDist.tab from RecombCalculatorDDose

Outputs:
A dotplot file representing pairwise marker linkage

History
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
50
51
52
53
54
55
56
57
58
59
60
62
63

1 (red arrow pointing to History panel)

2 (red arrow pointing to Image in png format button)

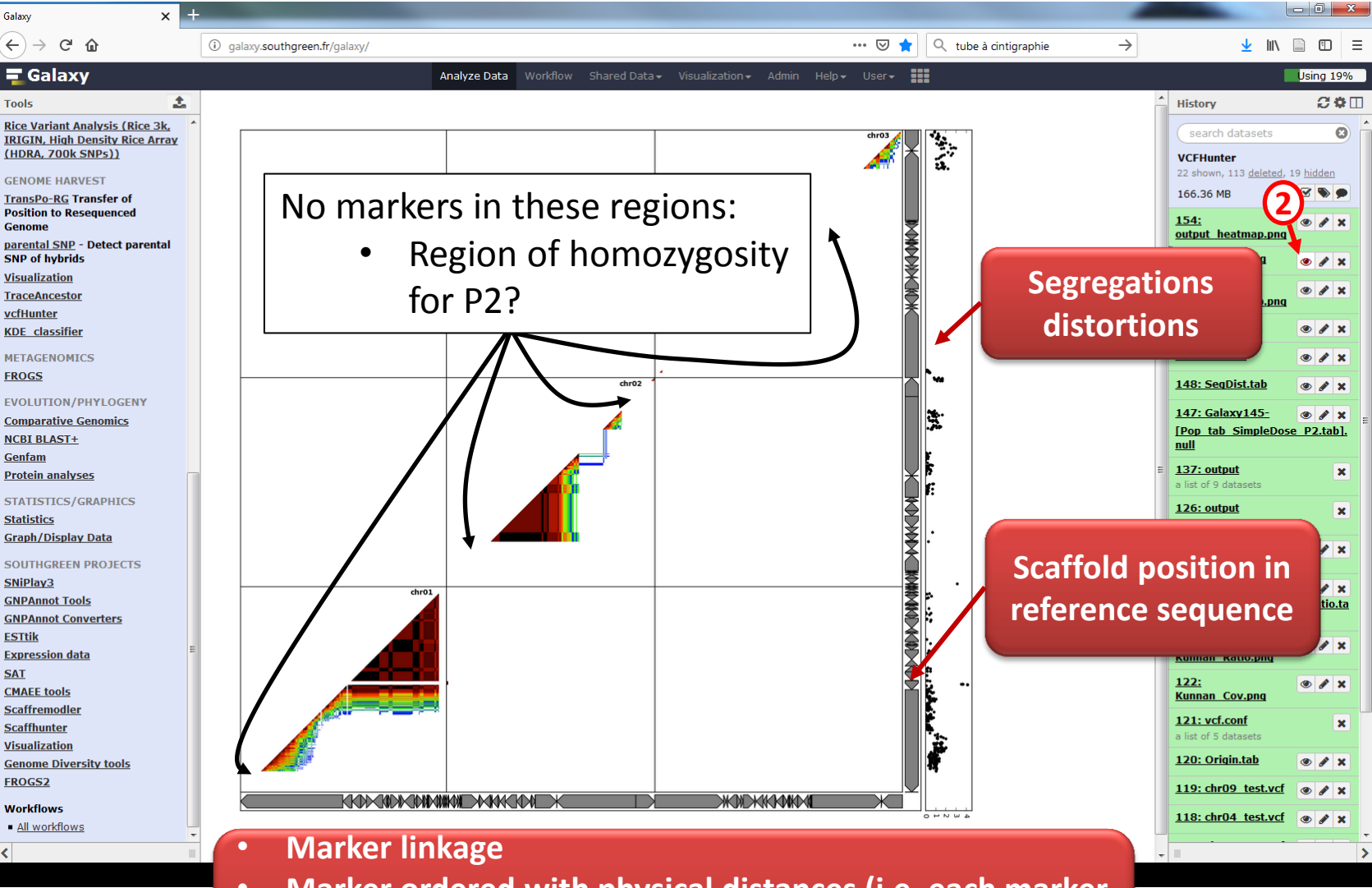
3 (red arrow pointing to CartoRef.agp)

4 (red arrow pointing to SegDist.tab)

5 (red arrow pointing to y)

Genetic mapping analysis with vcfHunter tool

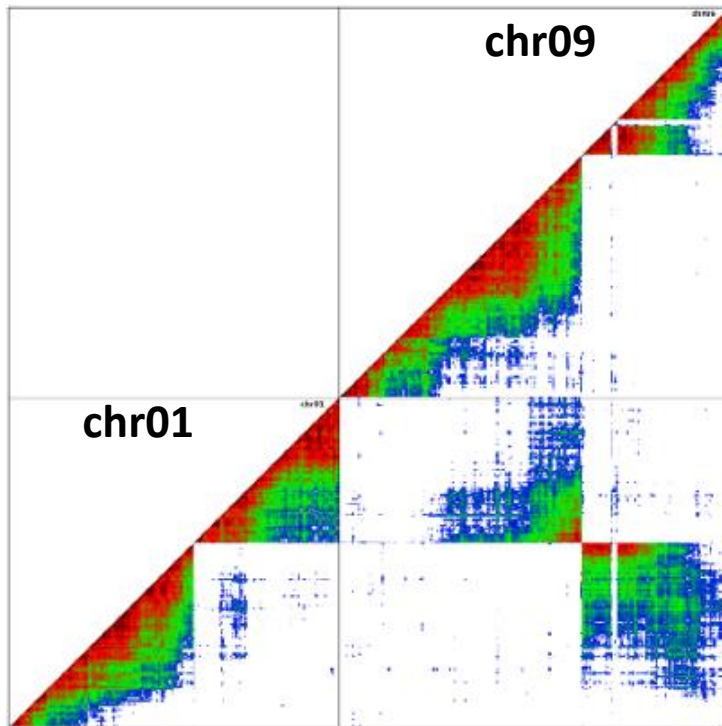
Plotting pairwise recombination with other options



Genetic mapping analysis with vcfHunter tool

Plotting pairwise recombination with other populations

Evidence of reciprocal translocations between chromosome 1 and 9



Evidence for large segregation distortions and inter-chromosomal linkage

