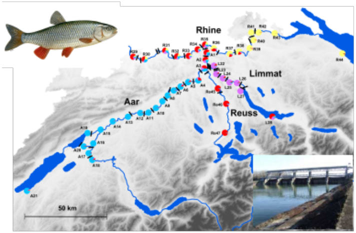
Analysis of genetic differentiation at the NGS era

Valentin Hivert

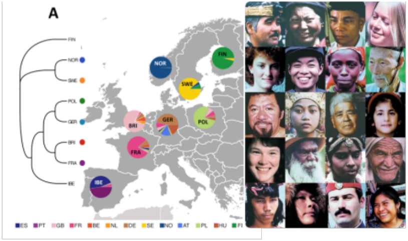
INRA CBGP Montferrier-sur-Lez

Thesis defense, December 14^{th} 2018

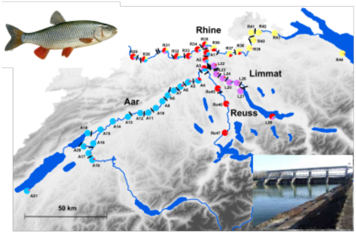
Supervisors : Renaud Vitalis, Mathieu Gautier



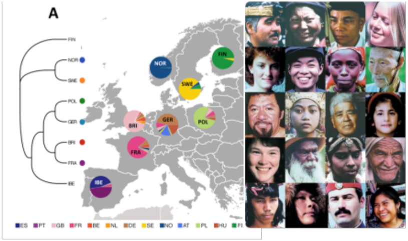
Gouskov & al. 2015



Athanasiadis & al. 2016



Gouskov & al. 2015



Athanasiadis & al. 2016

The spatial and temporal organisation of individuals in groups

(subpopulation, social group, family...) foster the genetic differentiation → differences in allele frequencies between groups

Evolutionary forces :

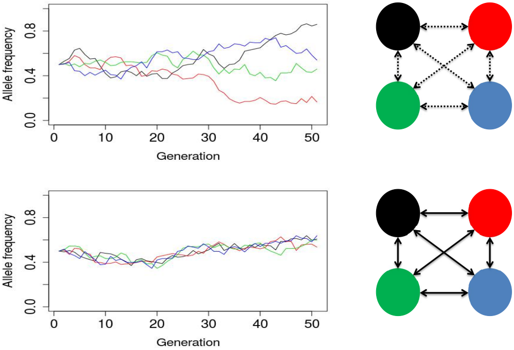
- Mutation
- Genetic drift
- Gene flow
- Selection

Evolutionary forces :

- Global effect :
  - Genetic drift
  - Gene flow
- Local effect :
  - Mutation
  - Selection

Effect of Gene flow and selection on genetic differentiation

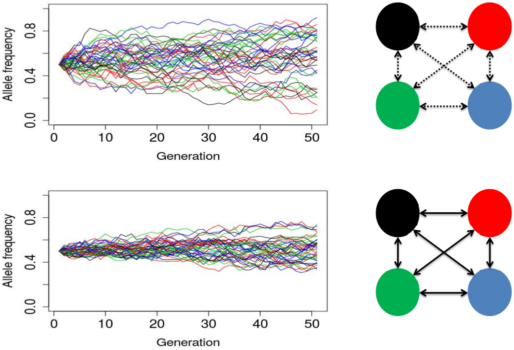
Genome-wide effect



- Homogenizes the allele frequencies → decreases the allele frequencies variance between demes

Effect of Gene flow and selection on genetic differentiation

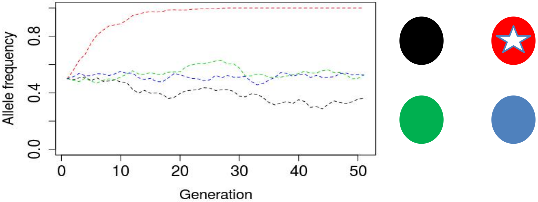
Genome-wide effect



- Homogenizes the allele frequencies $\rightarrow$ decreases the allele frequencies variance between demes

Effect of Gene flow and selection on genetic differentiation

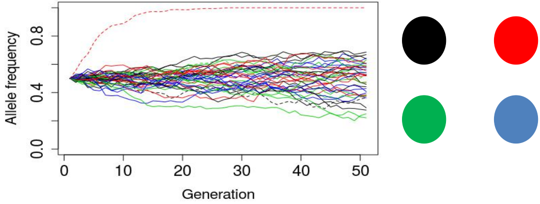
Local effect on the genome



- Increases the allele frequencies variance between demes

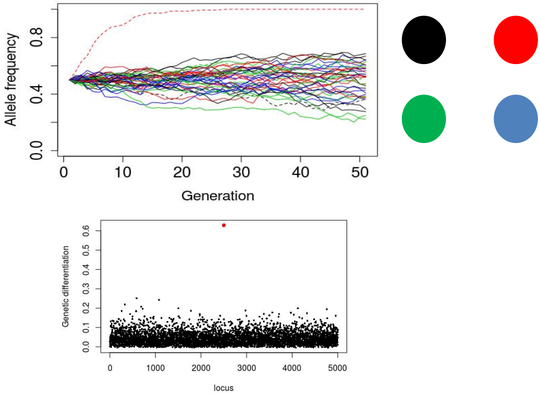
Effect of Gene flow and selection on genetic differentiation

Local effect on the genome



Effect of Gene flow and selection on genetic differentiation

Local effect on the genome



We need to characterize the genetic variability at a genomic scale

The genomic revolution

Next Generation Sequencing (NGS) :

- Very large numbers of markers
  $\rightarrow x10^6$ markers

The genomic revolution

Next Generation Sequencing (NGS) :

- Very large numbers of markers
  $\rightarrow x10^6$ markers
- Allows to characterize genetic variability at a pan-genomic scale and at a lower cost

The genomic revolution

Next Generation Sequencing (NGS) :

- Very large numbers of markers $\rightarrow x10^6$ markers
- Allows to characterize genetic variability at a pan-genomic scale and at a lower cost
- High density of markers allows the use of linkage information

The genomic revolution

Next Generation Sequencing (NGS) :

- Very large numbers of markers $\rightarrow x10^6$ markers
- Allows to characterize genetic variability at a pan-genomic scale and at a lower cost
- High density of markers allows the use of linkage information

NGS $\rightarrow$ change in the nature of data

Main research axis

My thesis focuses on the development of new statistical methods of genetic differentiation analysis from NGS data

- Development of an estimator of genetic differentiation, from NGS data

Main research axis

My thesis focuses on the development of new statistical methods of genetic differentiation analysis from NGS data

- Development of an estimator of genetic differentiation, from NGS data
- Development of a new method of genetic differentiation analysis, for the research of signature of selection from high density NGS data

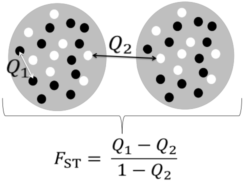
Part I : Measuring genetic differentiation from Pool-seq data

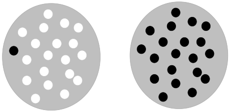
$F_{\mathrm{ST}} \to 0$ $F_{\mathrm{ST}} \to 1$

- F_{ST} is defined as the portion of the total genetic variance explained by the genetic variance between subpopulations

- F_{ST} is defined as the portion of the total genetic variance explained by the genetic variance between subpopulations
- F_{ST} is classically estimated under an analysis-of-variance framework (Weir & Cockerham 1984)

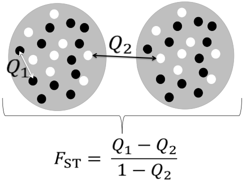
$F_{\mathrm{ST}} \to 0$ $F_{\mathrm{ST}} \to 1$



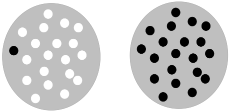


- It can be expressed in terms of probabilities of identity in states for pairs of genes (Cockerham 1973; Rousset 2007)

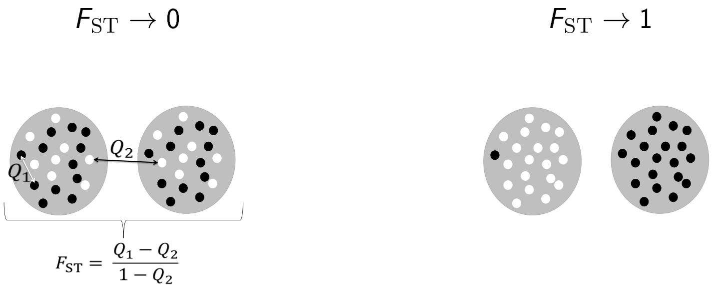
$F_{\mathrm{ST}} \to 0$



$F_{\mathrm{ST}} \to 1$



- It can be expressed in terms of probabilities of identity in states for pairs of genes (Cockerham 1973; Rousset 2007)
- F_{ST} can be estimated with $\hat{Q}_1$ and $\hat{Q}_2$



- It can be expressed in terms of probabilities of identity in states for pairs of genes (Cockerham 1973; Rousset 2007)
- F_{ST} can be estimated with $\hat{Q}_1$ and $\hat{Q}_2$

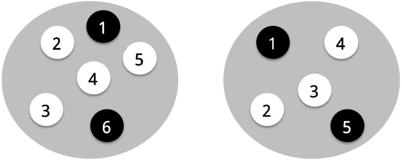
Equal sample sizes $\rightarrow$ strictly reduces to the analysis-of-variance estimator (Weir & Cockerham, 1984)

We are interested in the variance of allele frequencies at the population scale

The Pool-seq $\rightarrow$ a cost-effective alternative to individual genotyping

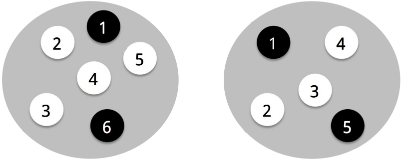
The Pool-seq process

pooling

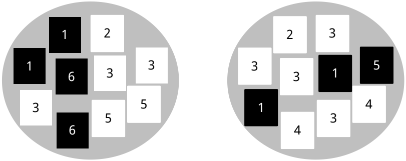


The Pool-seq process

pooling

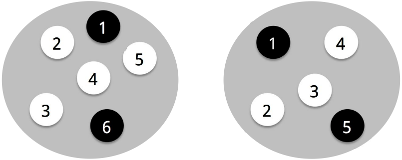


Sequencing (10x coverage)

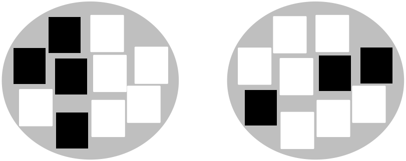


The Pool-seq process

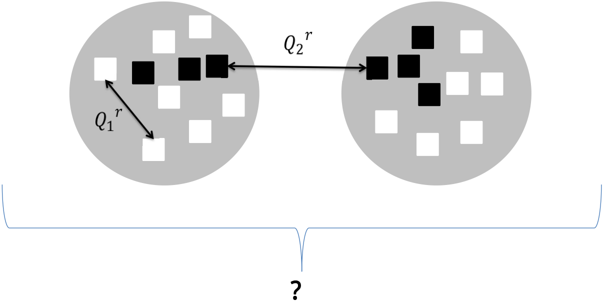
pooling



Sequencing (10x coverage)

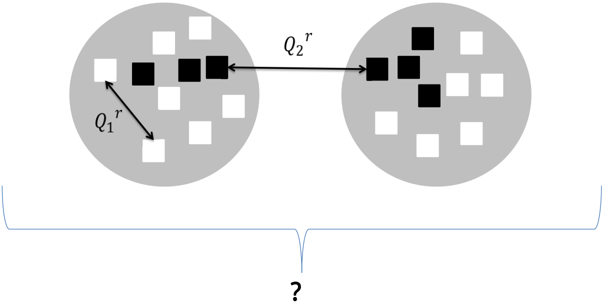


The Pool-seq process

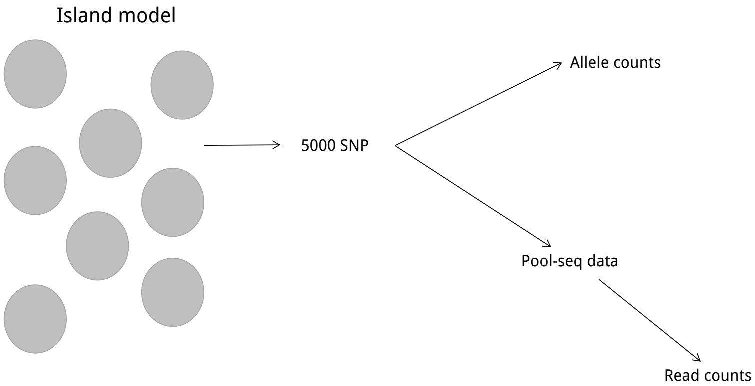


How can we estimate F_{ST} from Pool-seq data ?

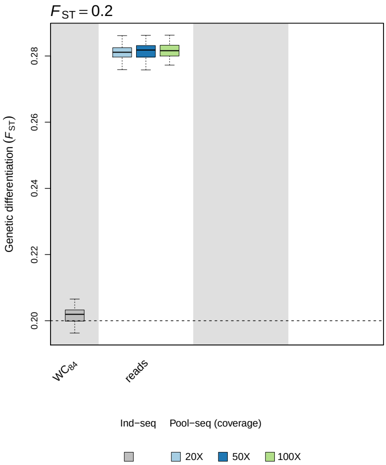
The Pool-seq process



$$\hat{F}_{\mathrm{ST}}^{reads} = \frac{\hat{Q}_1^r - \hat{Q}_2^r}{1 - \hat{Q}_2^r}$$

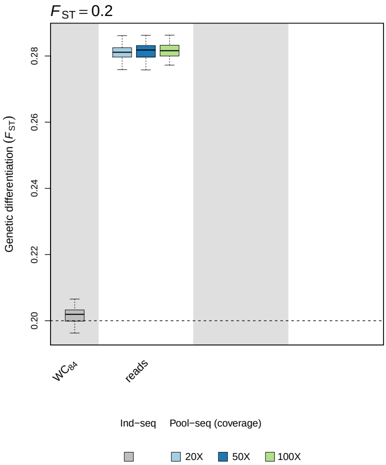

Island model
5000 SNP
Allele counts
Pool-seq data
Read counts

Island Model, $n_d = 8$, $N = 10$ and $F_{ST} = 0.2$



- WC_{84} : analysis-of-variance estimates (Weir & Cockerham 1984) computed from individual data (allele counts)
- reads : estimates computed directly from read counts IIS probabilities

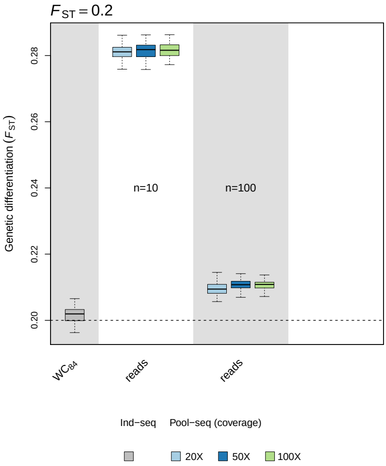
Island Model, $n_d = 8$, $N = 10$ and $F_{ST} = 0.2$



- WC_{84} : analysis-of-variance estimates (Weir & Cockerham 1984) computed from individual data (allele counts)
- reads : estimates computed directly from read counts IIS probabilities

Bias reads >> bias WC_{84}

Island Model, $n_d = 8$, $N = 10$ and $F_{\mathrm{ST}} = 0.2$

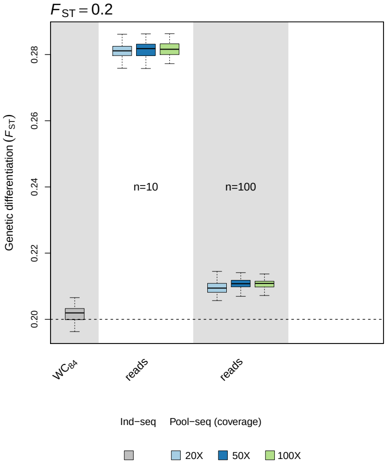


- WC_{84} : analysis-of-variance estimates (Weir & Cockerham 1984) computed from individual data (allele counts)
- reads : estimates computed directly from read counts IIS probabilities

Bias reads >> bias WC_{84}

The bias depends on the pool size

Island Model, $n_d = 8$, $N = 10$ and $F_{ST} = 0.2$



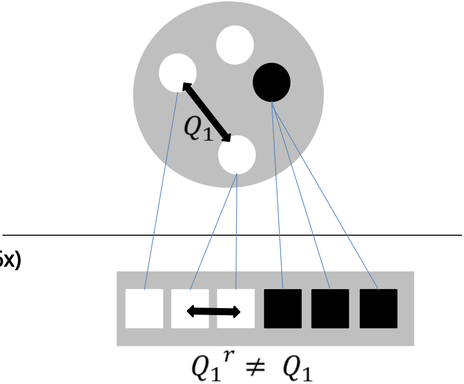
- WC_{84} : analysis-of-variance estimates (Weir & Cockerham 1984) computed from individual data (allele counts)
- reads : estimates computed directly from read counts IIS probabilities

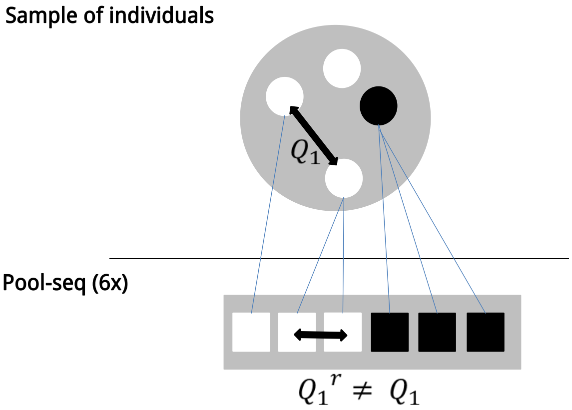
Bias reads >> bias WC_{84}

The bias depends on the pool size

Sample of individuals

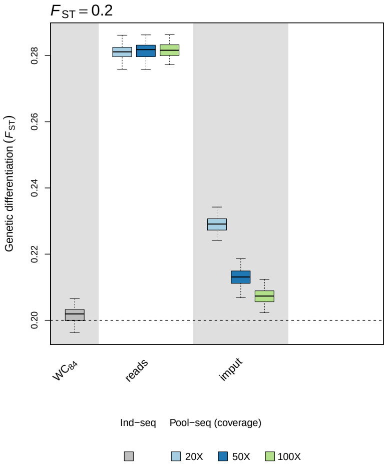
Pool-seq (6x)





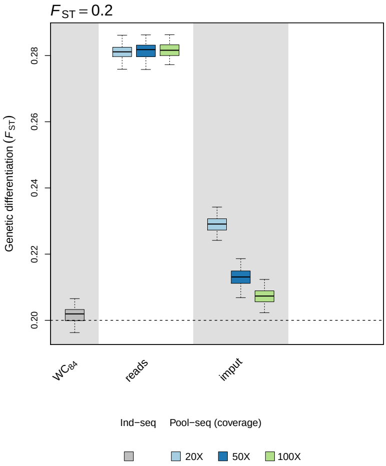
Alternative : estimation of individual counts by Maximum likelihood from reads frequencies and pool sizes

Island Model, $n_d = 8$, $N = 10$ and $F_{ST} = 0.2$



- imput : WC_{84} estimates computed from allele counts estimated by maximum-likelihood

Island Model, $n_d = 8$, $N = 10$ and $F_{ST} = 0.2$



- imput : WC_{84} estimates computed from allele counts estimated by maximum-likelihood

Bias Imput >> bias WC_{84}

The bias depends on the coverage

The model

We have developed $\hat{F}_{\mathrm{ST}}^{\mathrm{pool}}$, a new estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework[1]

- The total variance is decomposed into reads within individuals, individuals within demes and among demes

[1] Hivert et al. 2018.

The model

We have developed $\hat{F}_{\mathrm{ST}}^{\mathrm{pool}}$, a new estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework[1]

- The total variance is decomposed into reads within individuals, individuals within demes and among demes
- We assume an equal individual's contribution into the pool of DNA (multinomial distribution of the reads)

[1] Hivert et al. 2018.

The model

We have developed $\hat{F}_{\mathrm{ST}}^{\mathrm{pool}}$, a new estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework[1]

- The total variance is decomposed into reads within individuals, individuals within demes and among demes
- We assume an equal individual's contribution into the pool of DNA (multinomial distribution of the reads)

$$\hat{F}_{\mathrm{ST}}^{\mathrm{pool}} = \frac{\sum_k \left[(C_1 - D_2) \sum_i^{n_{\mathrm{d}}} C_{1i} (\hat{\pi}_{i:k} - \hat{\pi}_k)^2 - (D_2 - D_2^\star) \sum_i^{n_{\mathrm{d}}} C_{1i} \hat{\pi}_{i:k} (1 - \hat{\pi}_{i:k}) \right]}{\sum_k \left[(C_1 - D_2) \sum_i^{n_{\mathrm{d}}} C_{1i} (\hat{\pi}_{i:k} - \hat{\pi}_k)^2 + (n_{\mathrm{c}} - 1) \left(D_2 - D_2^\star \right) \sum_i^{n_{\mathrm{d}}} C_{1i} \hat{\pi}_{i:k} (1 - \hat{\pi}_{i:k}) \right]}$$

[1] Hivert et al. 2018.

The model

We have developed $\hat{F}_{\mathrm{ST}}^{\mathrm{pool}}$, a new estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework[1]

- The total variance is decomposed into reads within individuals, individuals within demes and among demes
- We assume an equal individual's contribution into the pool of DNA (multinomial distribution of the reads)

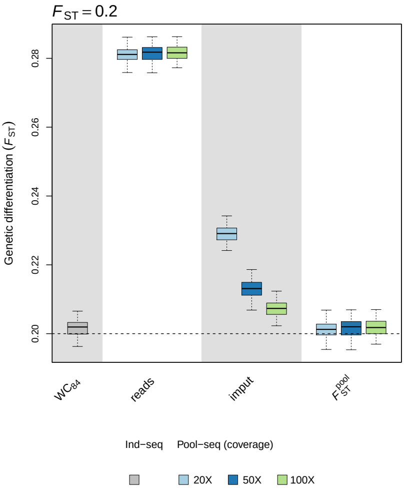
$$\hat{F}_{\mathrm{ST}}^{\mathrm{pool}} = \frac{\sum_k \left[(C_1 - D_2) \sum_i^{n_{\mathrm{d}}} C_{1i} (\hat{\pi}_{i:k} - \hat{\pi}_k)^2 - (D_2 - D_2^\star) \sum_i^{n_{\mathrm{d}}} C_{1i} \hat{\pi}_{i:k} (1 - \hat{\pi}_{i:k}) \right]}{\sum_k \left[(C_1 - D_2) \sum_i^{n_{\mathrm{d}}} C_{1i} (\hat{\pi}_{i:k} - \hat{\pi}_k)^2 + (n_{\mathrm{c}} - 1) \left(D_2 - D_2^\star \right) \sum_i^{n_{\mathrm{d}}} C_{1i} \hat{\pi}_{i:k} (1 - \hat{\pi}_{i:k}) \right]}$$

- We show that, in the limit case where all pools have the same size n:

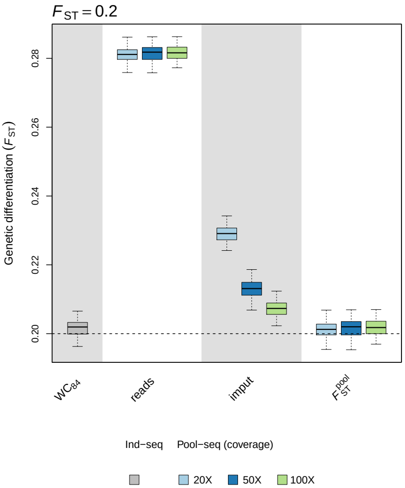
$$\hat{F}_{\mathrm{ST}}^{\mathrm{pool}} = 1 - \left(\frac{1 - \hat{Q}_1^{\mathrm{r}}}{1 - \hat{Q}_2^{\mathrm{r}}} \right) \left(\frac{n}{n - 1} \right)$$

[1] Hivert et al. 2018.

Island Model, $n_d = 8$, $N = 10$ and $F_{ST} = 0.2$



Island Model, $n_d = 8$, $N = 10$ and $F_{\mathrm{ST}} = 0.2$



Bias $\hat{F}_{\mathrm{ST}}^{\mathrm{pool}} \simeq$ bias WC_{84}

Independently on pool size, coverage and F_{ST} value

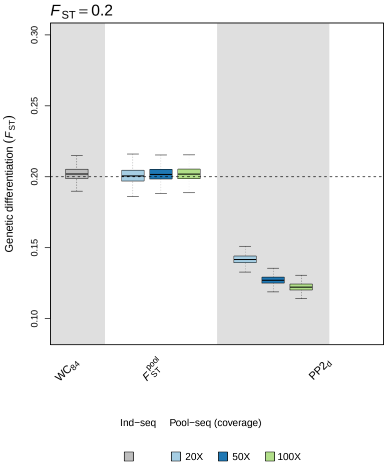
BIOINFORMATICS APPLICATIONS NOTE
Vol. 27 no. 24 2011, pages 3435–3436
doi:10.1093/bioinformatics/btr589

Genetics and population analysis
Advance Access publication October 23, 2011

PoPoolation2: identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq)

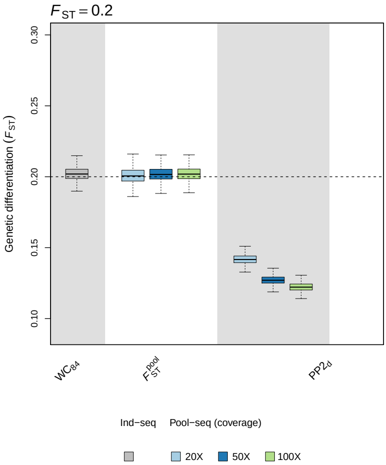
Robert Kofler, Ram Vinay Pandey and Christian Schlötterer*
Institut für Populationsgenetik, Vetmeduni Vienna, Veterinärplatz 1, A-1210 Wien, Austria
Associate Editor: Jeffrey Barrett

Island Model, $n_d = 8$, $N = 100$ and $F_{ST} = 0.2$



- $\mathrm{PP2_d}$: Popoolation2 estimator computed from read counts

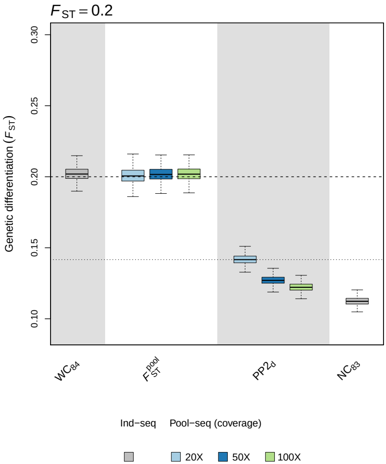
Island Model, $n_d = 8$, $N = 100$ and $F_{ST} = 0.2$



- $\mathrm{PP2_d}$: Popoolation2 estimator computed from read counts

$\mathrm{PP2_d}$ estimates are biased and it depends on the coverage.

Island Model, $n_d = 8$, $N = 100$ and $F_{ST} = 0.2$



- NC_{83} : Heterozygosity based estimator (Nei & Chesser 1983) computed from individual data
- $PP2_d$: Popoolation2 estimator computed from read counts

$PP2_d$ estimates are biased and it depends on the coverage.
It converges to the Nei and Chesser's estimator $(NC_{83})^2$ as the coverage increases.

[2] Nei and Chesser 1938.

MOLECULAR ECOLOGY

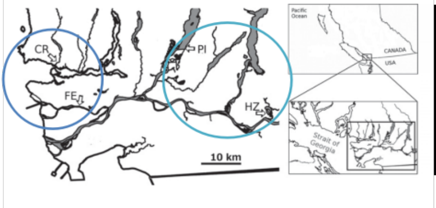
Molecular Ecology (2017) **26**, 25–42 doi: 10.1111/mec.13805

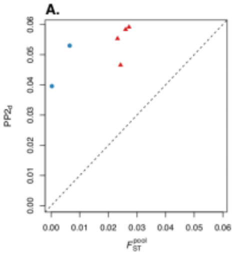
SPECIAL ISSUE: THE MOLECULAR MECHANISMS OF ADAPTATION AND SPECIATION: INTEGRATING GENOMIC AND MOLECULAR APPROACHES

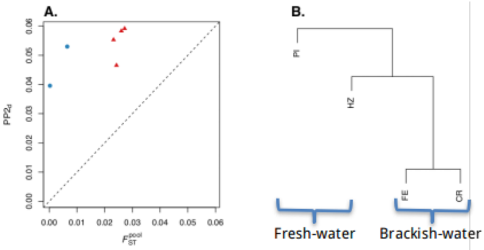
Adaptive genomic divergence under high gene flow between freshwater and brackish-water ecotypes of prickly sculpin (*Cottus asper*) revealed by Pool-Seq

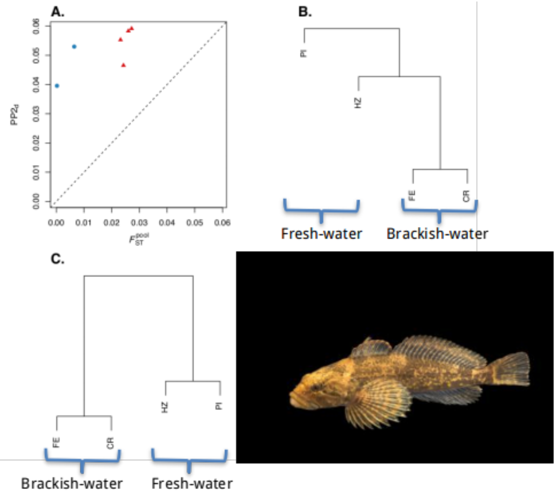
STEFAN DENNENMOSER,*† STEVEN M. VAMOSI,† ARNE W. NOLTE*‡ and SEAN M. ROGERS†
**Max-Planck Institute for Evolutionary Biology, August Thienemann Strasse 2, 24306, Plön, Germany, †Department of Biological Sciences, University of Calgary, 2500 University Drive NW, Calgary AB, Canada T2N 1N4, ‡Institute for Biology, Carl von Ossietzky University Oldenburg, Carl von Ossietzky Str. 9-11, 26111 Oldenburg, Germany*

Brackish-water Fresh-water




A.
$PP2_d$
0.00
0.01
0.02
0.03
0.04
0.05
0.06
F_{ST}^{pool}


A.
PP2d
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0.00
0.01
0.02
0.03
0.04
0.05
0.06
F_{ST}^{pool}
B.
PI
HZ
FE
CR
Fresh-water
Brackish-water


A.
B.
C.
$PP2_d$
F_{ST}^{pool}
PI
HZ
FE
CR
Fresh-water
Brackish-water

Conclusion

We developed an unbiased estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework.

- The accuracy is barely distinguishable from the analysis-of-variance estimator for individual data (Weir & Cockerham, 1984).

Conclusion

We developed an unbiased estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework.

- The accuracy is barely distinguishable from the analysis-of-variance estimator for individual data (Weir & Cockerham, 1984).
- The accuracy does not depend on the coverage or on the pool size.

Conclusion

We developed an unbiased estimator of F_{ST} for Pool-seq data, in an analysis-of-variance framework.

- The accuracy is barely distinguishable from the analysis-of-variance estimator for individual data (Weir & Cockerham, 1984).
- The accuracy does not depend on the coverage or on the pool size.
- Although our estimator is sensitive to uneven contributions of individual DNAs in each pool, we found that it was robust to sequencing errors, ascertainment bias, unequal sample sizes and variable coverages.

Conclusion

- We focused on global (multi-locus) genetic differentiation

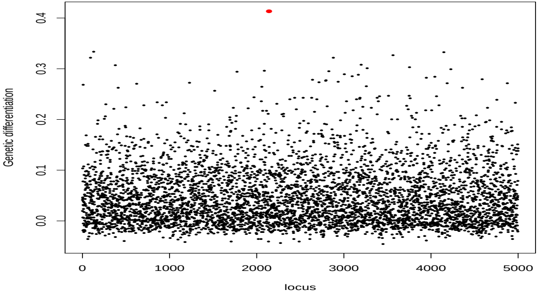
Conclusion

- We focused on global (multi-locus) genetic differentiation

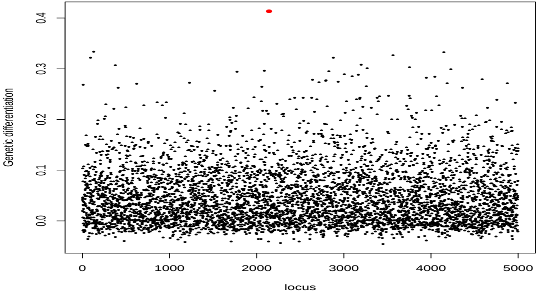
What about selection ?

- It has been proposed to identify loci under selection from genomic scan of differentiation

Conclusion

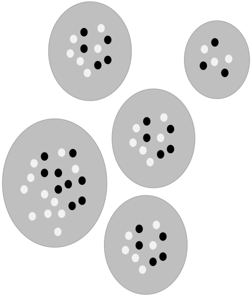

Genetic differentiation
0.0
0.1
0.2
0.3
0.4
0
1000
2000
3000
4000
5000
locus

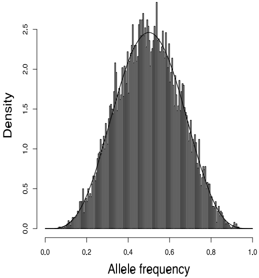
Conclusion



- How to distinguish local effect (selection) from global effect (demography) ?

Part II : A hierarchical Bayesian model for measuring the extent of local adaptation using linkage disequilibrium information





Allele frequencies distribution can be characterized conditionally on some demo-genetic model

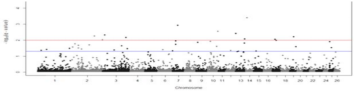
DOI: 10.1534/genetics.108.092221

A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective

Matthieu Foll[1] and Oscar Gaggiotti

Laboratoire d'Ecologie Alpine (LECA), CNRS UMR 5553, 38041 Grenoble Cedex 09, France

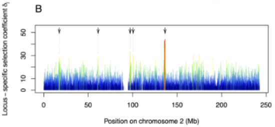
Manuscript received June 3, 2008
Accepted for publication July 23, 2008



INVESTIGATION
HIGHLIGHTED ARTICLE

Detecting and Measuring Selection from Gene Frequency Data

Renaud Vitalis,*,†,1 Mathieu Gautier,*,† Kevin J. Dawson,‡ and Mark A. Beaumont§
*Institut National de la Recherche Agronomique, Unité Mixte de Recherche CBGP, (Inra, Ird, Cirad, Montpellier-SupAgro) 34988 Montferrier-sur-Lez Cedex, France, †Institut de Biologie Computationnelle, 34095 Montpellier Cedex, France, ‡Cancer Genome ...ust Sanger Institute, Hinxton, CB10 1SA, United Kingdom, §Department of Mathematics and School of Biological Sciences, University of Bristol, Bristol BS8 1TW, United Kingdom



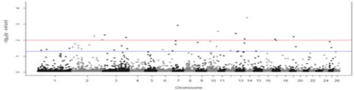
DOI: 10.1534/genetics.108.092221

A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective

Matthieu Foll[1] and Oscar Gaggiotti

Laboratoire d'Ecologie Alpine (LECA), CNRS UMR 5553, 38041 Grenoble Cedex 09, France

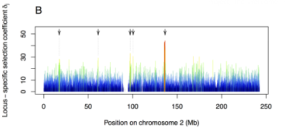
Manuscript received June 3, 2008
Accepted for publication July 23, 2008



INVESTIGATION
HIGHLIGHTED ARTICLE

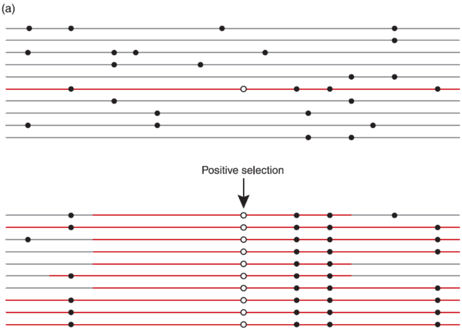
Detecting and Measuring Selection from Gene Frequency Data

Renaud Vitalis,*,†,1 Mathieu Gautier,*,† Kevin J. Dawson,‡ and Mark A. Beaumont§
*Institut National de la Recherche Agronomique, Unité Mixte de Recherche CBGP, (Inra, Ird, Cirad, Montpellier-SupAgro) 34988 Montferrier-sur-Lez Cedex, France, †Institut de Biologie Computationnelle, 34095 Montpellier Cedex, France, ‡Cancer Genome ...ust Sanger Institute, Hinxton, CB10 1SA, United Kingdom, §Department of Mathematics and School of Biological Sciences, University of Bristol, Bristol BS8 1TW, United Kingdom



Most methods generally neglect the information brought by linkage disequilibrium (LD) among genetic markers

Hard-sweep



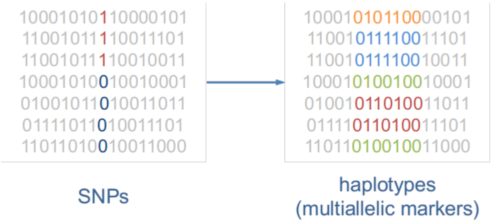
3

[3] Storz 2005.

How to account for LD information?

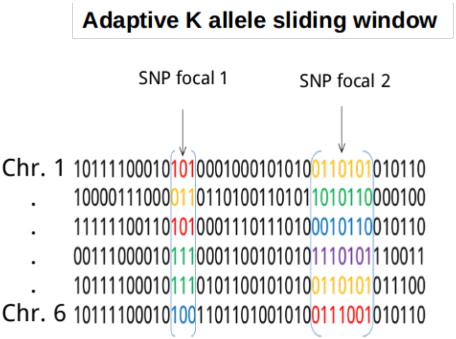
How to account for LD information?

→ Extend SelEstim (Vitalis et al. 2014), a hierarchical bayesian model to the use of multi-allelic markers



How to account for LD information?

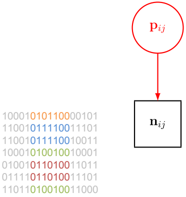
→ Extend SelEstim (Vitalis et al. 2014), a hierarchical bayesian model to the use of multi-allelic markers



The model

The data : haplotypes at many loci, in several populations (allele counts)

The model

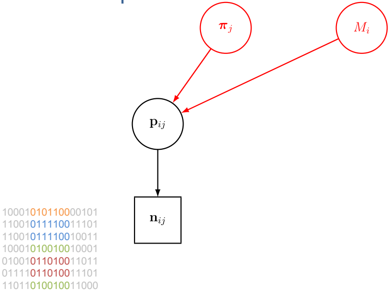


The (unknown) allele frequencies. Approximation of a diffusion process as prior distribution

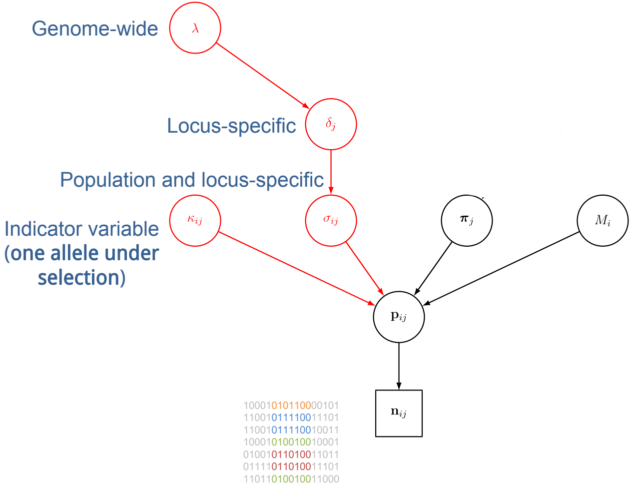
➔ migration-drift-selection equilibrium

The model

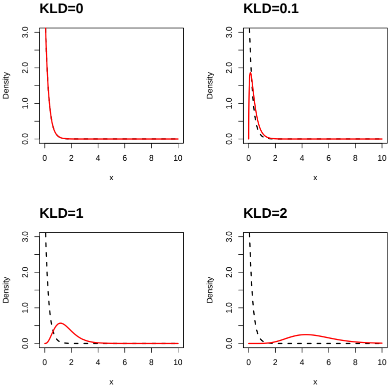
Infinite island model: the population frequencies depend on $M_i = 4N_i m_i$ and the frequencies in the migrant pool



The model


Genome-wide
λ
Locus-specific
δ_j
Population and locus-specific
κ_{ij}
σ_{ij}
$\boldsymbol{\pi}_j$
M_i
Indicator variable
(one allele under selection)
$\mathbf{p}_{ij}$
$\mathbf{n}_{ij}$
10001010110000101
11001011110011101
11001011110010011
10001010010010001
01001011010011011
01111011010011101
11011010010011000

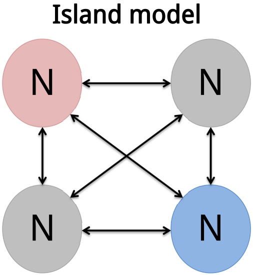
The decision criterion



- We use the Kullback-Leibler Divergence (KLD) as a distance between the posterior distributions of the δ_j's and a centering distribution

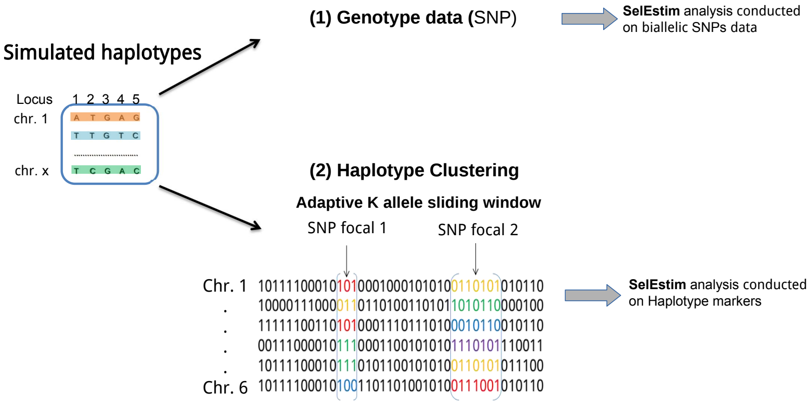
Evaluation by simulations

individual-based forward-time simulations with demography and selection

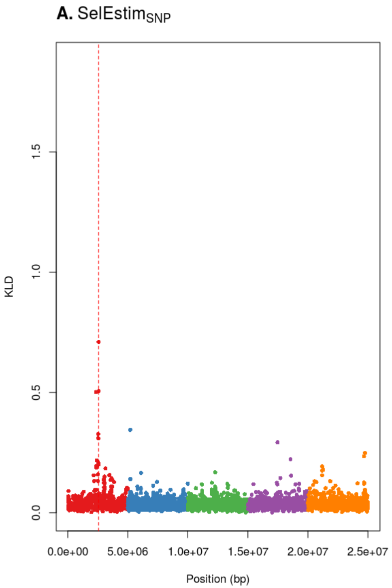


N = 1000 diploid individuals
5 chromosomes of 5 Mb (selection on chromosome 1)
density of markers : 125 SNP/Mb
500 replicates per scenario

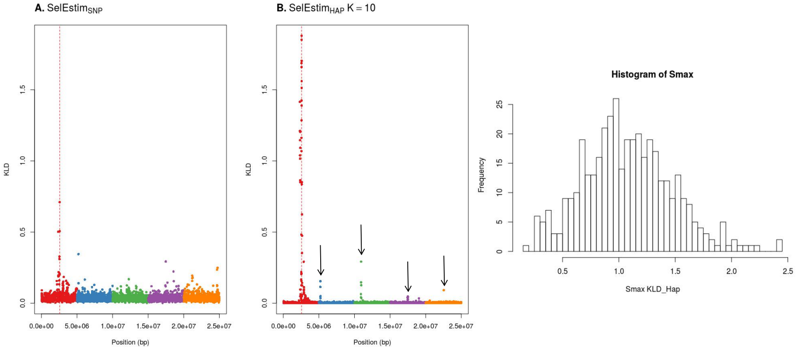
Evaluation by simulations



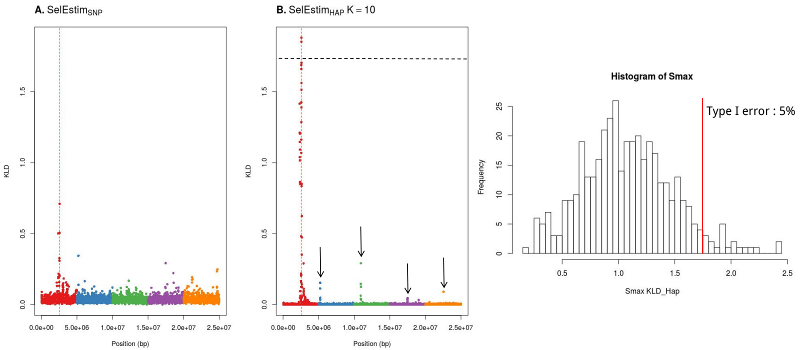
Example of SelEstim outputs



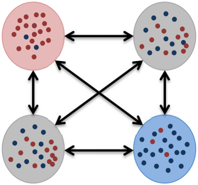
Method of analysis

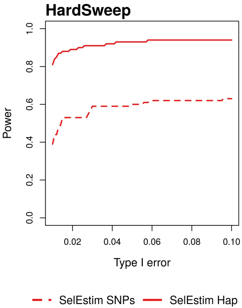


Method of analysis



Power for Island Model





- Improved statistical power with haplotype-based analyses (*vs.* SNPs)

Power for Island Model

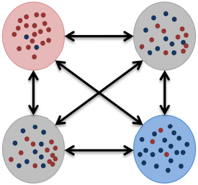
- FLK[4] is an extent of the LK test (Lewontin and Krakauer 1973) to account for the hierarchical structure of populations
- HapFLK[5] extent the model FLK to the use of haplotype data (HapFLK has is own clustering algorithm)

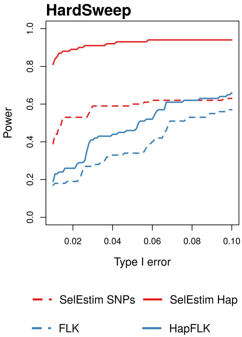
Both models are expected to better perform under a pure drift demography

[4] Bonhomme et al. 2010.

[5] Fariello et al. 2013.

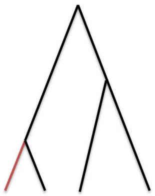
Power for Island Model

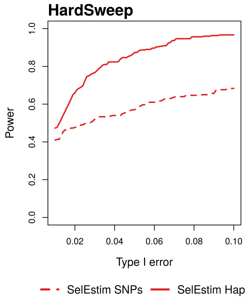




- Improved statistical power with haplotype-based analyses (*vs.* SNPs)
- Outperform FLK and HapFLK

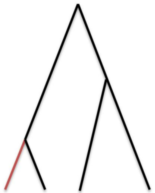
Power for Pure Drift Model

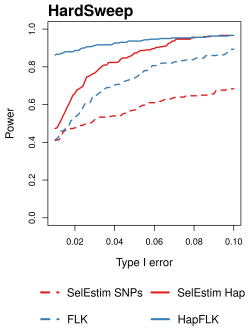




- Improved statistical power with haplotype-based analyses (*vs.* SNPs)

Power for Pure Drift Model

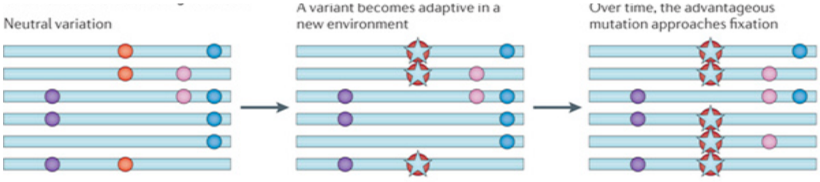




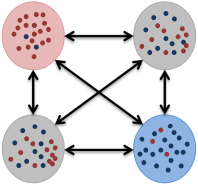
- Improved statistical power with haplotype-based analyses (*vs.* SNPs)
- Fall behind FLK and HapFLK

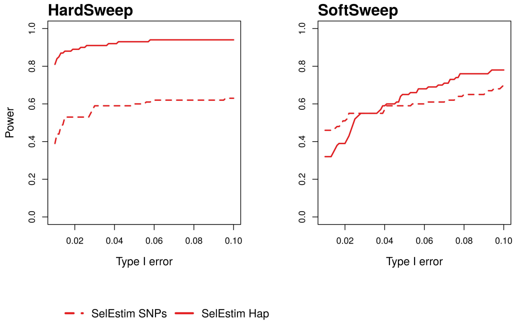
We considered hard-sweep scenarios. What happens with soft-sweep?

We considered hard-sweep scenarios. What happens with soft-sweep?

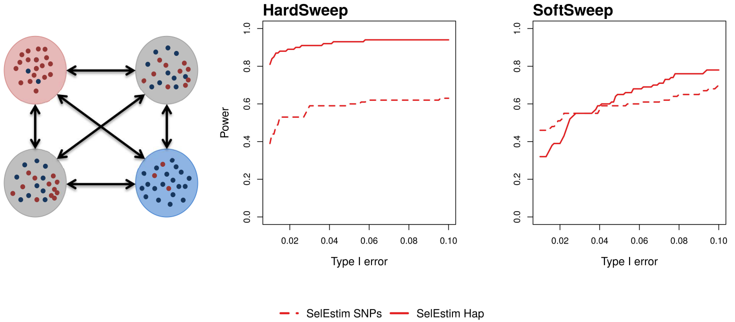


Power for Island Model with Soft sweep





Power for Island Model with Soft sweep



Soft-sweep → many alleles under selection (departure from the model assumption)

Conclusion

We developed a hierarchical bayesian model to measure the extent of local adaptation from haplotype data.

- LD information brought by haplotype data $\rightarrow$ Increases the detection power of selection

Conclusion

We developed a hierarchical bayesian model to measure the extent of local adaptation from haplotype data.

- LD information brought by haplotype data → Increases the detection power of selection
- Be aware of the underlying demo-genetic models and assumptions as well as the robustness of the methods to model misspecifications

General conclusion and perspectives

In this thesis, I developed new statistical methods of genetic differentiation analysis for NGS data in different framework :

A summary statistic of F_{ST} for Pool-seq data in a frequentist approach

- To properly estimate the genetic differentiation from Pool-seq data, we need to account for the different levels of sampling
- Use of biased estimators $\rightarrow$ problem for genome scan when variable coverage on the genome

General conclusion and perspectives

In this thesis, I developped new statistical methods of genetic differentiation analysis for NGS data in different framework :

A hierarchical bayesian model for the detection of signature of selection from haplotype data

- LD information brought by high density data increases the power to detect selection
- We considered an equilibrium model $\rightarrow$ beware of confonding effects (allele surfing...)

General conclusion and perspectives

In this thesis, I developped new statistical methods of genetic differentiation analysis for NGS data in different framework :

A hierarchical bayesian model for the detection of signature of selection from haplotype data

- LD information brought by high density data increases the power to detect selection
- We considered an equilibrium model → beware of confonding effects (allele surfing...)

The nature of the data used in the two parts are different

General conclusion and perspectives

Is it possible to estimate haplotype frequencies from Pool-seq ?

- Models exist but need information about the pool of haplotypes (Cao et Sun 2015; Kessner et al. 2013; Long et al. 2011) or are specifically designed for E&R experiences (Franssen et al. 2017).

General conclusion and perspectives

Is it possible to estimate haplotype frequencies from Pool-seq ?

- Models exist but need information about the pool of haplotypes (Cao et Sun 2015; Kessner et al. 2013; Long et al. 2011) or are specifically designed for E&R experiences (Franssen et al. 2017).

Is it possible to account for LD with unphased data (i.e Pool-seq) ?

- Investigation of a smoothing model incorporate in SelEstim to account for the spatial correlation between markers

General conclusion and perspectives

Genome scans are a first step to identifying putative genomic regions under selection

- Poor reproducibility among methods (Pritchard et al. 2010)
- Functional validation of candidate genes

Acknowledgments

The jury members

- Christine Dillmann (R)
- Anna-Sapfo Malaspinas (R)
- Miguel Pérez-Enciso (E)
- Joëlle Ronfort (E)

My supervisors

- Renaud Vitalis
- Mathieu Gautier

The comitees members

- Stephanie Manel
- Michael Blum
- Simon Boitard
- Bertrand Servin

The "Team" colleagues

- Arnaud Estoup
- Raphaël Leblois
- Miguel Navascués
- Alexandre Dehne-Garcia

Thanks to all of the CBGP colleagues and friends