

La variation clonale chez *Vitis vinifera* L.

Carrier Grégory

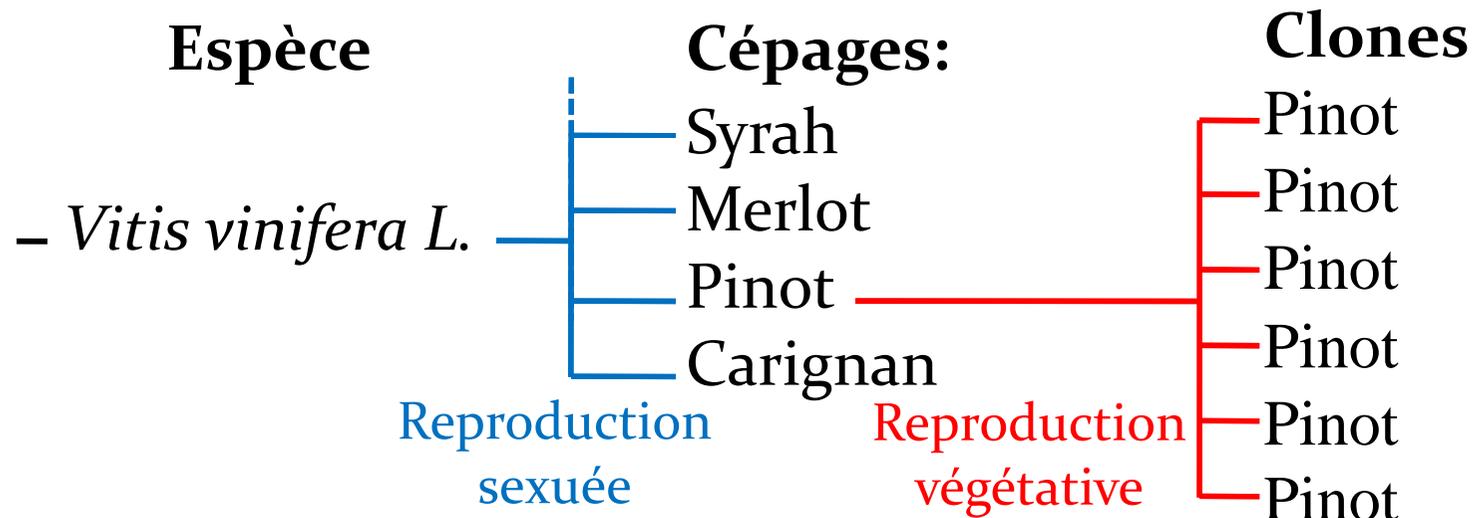
Directeurs de thèse: Patrice This
Jean-Michel Boursiquot

Encadrant: Loïc Le Cunff



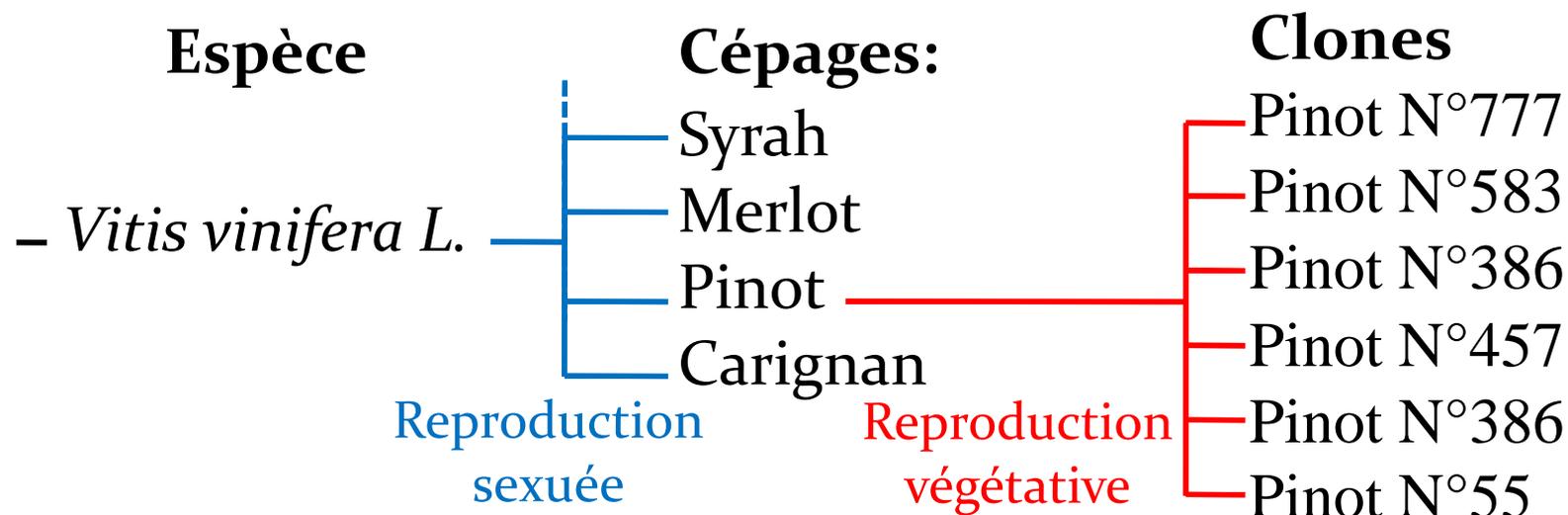
Un clone c'est...

Individus → Obtenus par reproduction végétative
à partir d'un cépage unique



Un clone c'est...

- Individus → Obtenus par reproduction végétative
à partir d'un cépage unique
- Caractéristiques phénotypiques propres.



Objectif: mieux comprendre l'origine de la diversité clonale chez la vigne:

Stratégie:

**Identification et quantification sans *a priori* de(s)
mutation(s) majoritaire(s)**

Objectif: mieux comprendre l'origine de la diversité clonale chez la vigne:

Stratégie:

Identification et quantification sans *a priori* de(s)
mutation(s) majoritaire(s)

1-Séquençage du génome des clones en NGS

Clone 1



```
VTTTCTTGTCATTCCCAGTACTGCTGCCTGATTTTGAATGCTCCCTGAAGAAGCTAACATTCTTGTAGGAAAATATTGGACA  
TATTCCACCATTTTTGAGTATAGAGATGCATCTACATTTTTGTAGTTTTCTTTATAGGCCAAGAGAGAGTCCCTTAAGGA  
GAAAACTCCCTTTTATAACCACTTTTCTTGTGAATGAGCATAAGTAAGTGATTTTCGTAGTAGGTCAACAAACTTTTTGCA  
ATAAGGATAGTTGGACTCGTCTTTTTATACACAAGTTAGTCGCTTTTTTTATGGTTCTGGGTTTGGGATGAGATGATTTTAT
```

Clone 2



```
TATTCTTGTCATTCCCAGTACTGCTGCCTGATTTTGAATGCTCCCTGAAGAAGCTAACATTCTTGTAGGAAAATATTGGACA  
TATTCCACCATTTTTGAGTATAGAGATGCATCTACATTTTTGTAGTTTTCTTTATAGGCCAAGAGAGAGTCCCTTAAGGA  
AGAAAACTCCCTTTTATAACCACTTTTCTTGTGAATGAGCATAAGTAAGTGATTTTCGTAGTAGGTCAACAAACTTTTTGCA  
BATAAGGATAGTTGGACTCGTCTTTTTATACACAAGTTAGTCGCTTTTTTTATGGTTCTGGGTTTGGGATGAGATGATTTTAT
```

Objectif: mieux comprendre l'origine de la diversité clonale chez la vigne:

Stratégie:

Identification et quantification sans *a priori* de(s)
mutation(s) majoritaire(s)

- 1-Séquençage du génome des clones en NGS
- 2-Comparaison de la séquence de différents génomes

Clone 1

TTGTTTTGATTTCTTTAAATCATCTTATATATAGCGCTAGCTACGTACGATCGTCGCTGCCGGCTATATAA*****TTGCCGTCGCGGGCGGCGGTAT
TTGTTTTGATTTCTTTAAATCATCTTATATATAGCGCTAGCTACGTACGATCGTCGCTGCCGGCTATATAA*****TTGCCGTCGCGGGCGGCGGTAT

Clone 2

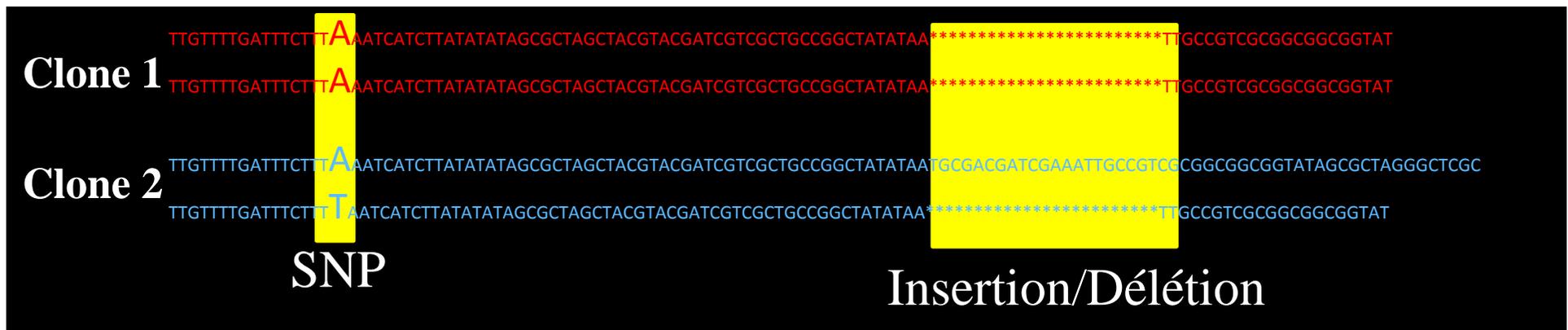
TTGTTTTGATTTCTTTAAATCATCTTATATATAGCGCTAGCTACGTACGATCGTCGCTGCCGGCTATATAATGCGACGATCGAAATTGCCGTCGCGGGCGGCGGTATAGCGCTAGGGCTCGC
TTGTTTTGATTTCTTTAAATCATCTTATATATAGCGCTAGCTACGTACGATCGTCGCTGCCGGCTATATAA*****TTGCCGTCGCGGGCGGCGGTAT

Objectif: mieux comprendre l'origine de la diversité clonale chez la vigne:

Stratégie:

Identification et quantification sans *a priori* de(s) mutation(s) majoritaire(s)

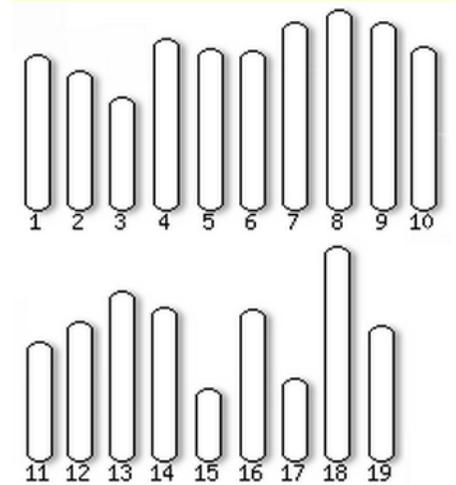
- 1-Séquençage du génome des clones en NGS
- 2-Comparaison de la séquence de différents génomes
- 3-Recherche de polymorphismes génétiques



Le génome de la vigne

19 chromosomes

470 Mb, ~ 6 fois plus petit que le génome humain



Séquences disponibles:

ATGCCGATGCATCGCAT
CTAGCATCGATCGATCA
CTAGCATCGATCGATCG
CGATGCATCAGCTA
GATCGATCGATCGATCA
GCATGCTACATCGATCA

l'individu PN 40024 (12X) (Jaillon et al., 2007)



ATGCCGATGCATCGCAT
CTAGCATCGATCGATCA
CTAGCATCGATCGATCG
CGATGCATCAGCTA
GATCGATCGATCGATCA
GCATGCTACATCGATCA

Le Pinot clone n° 115 (6,4X) (Vélasco et al., 2007)



Stratégie générale

Différents clones



Extraction d'ADN pour NGS

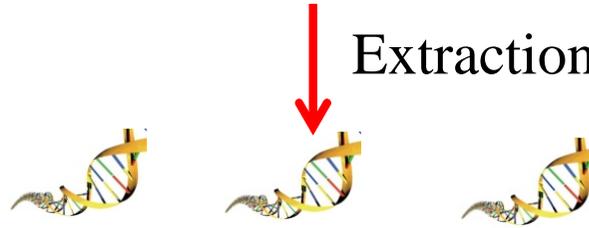


Stratégie générale

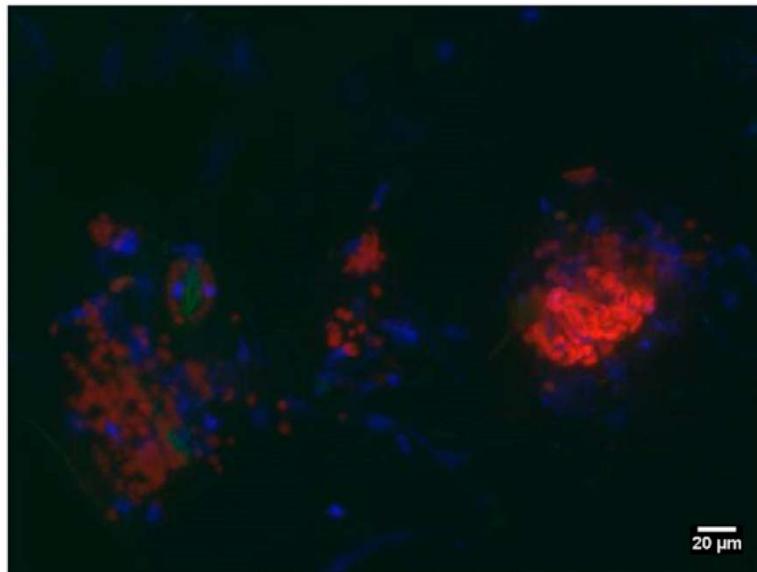
Différents clones



Extraction d'ADN pour NGS



Protocole classique



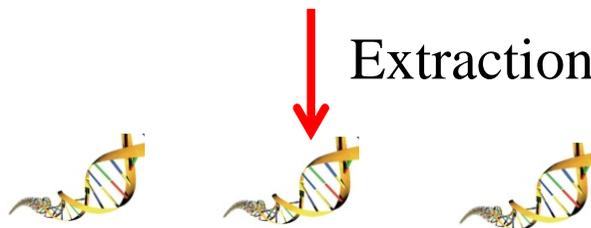
total DNA, Vitis

Stratégie générale

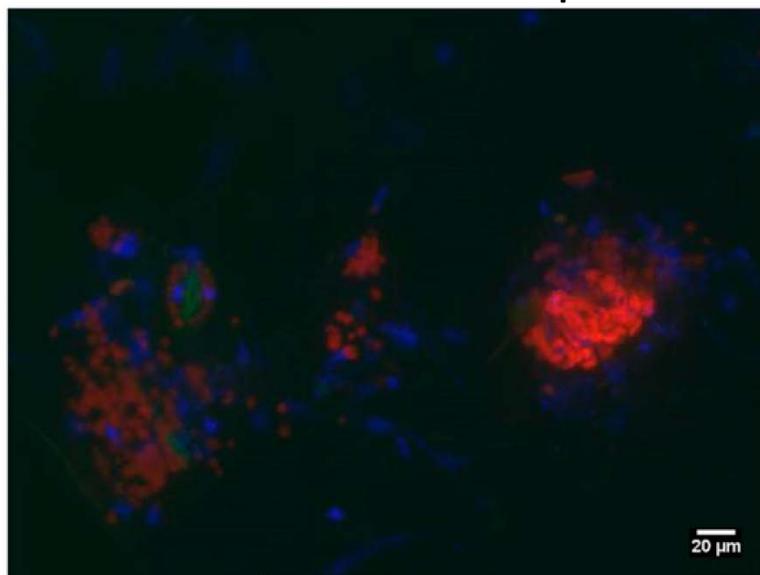
Différents clones



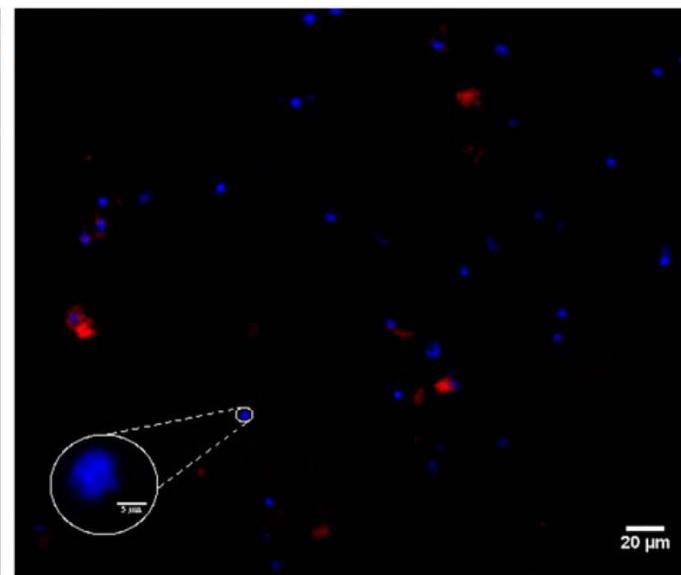
Extraction d'ADN pour NGS



Protocole classique



total DNA, Vitis



nuclear DNA, Vitis

Stratégie générale

Différents clones



Extraction d'ADN pour NGS



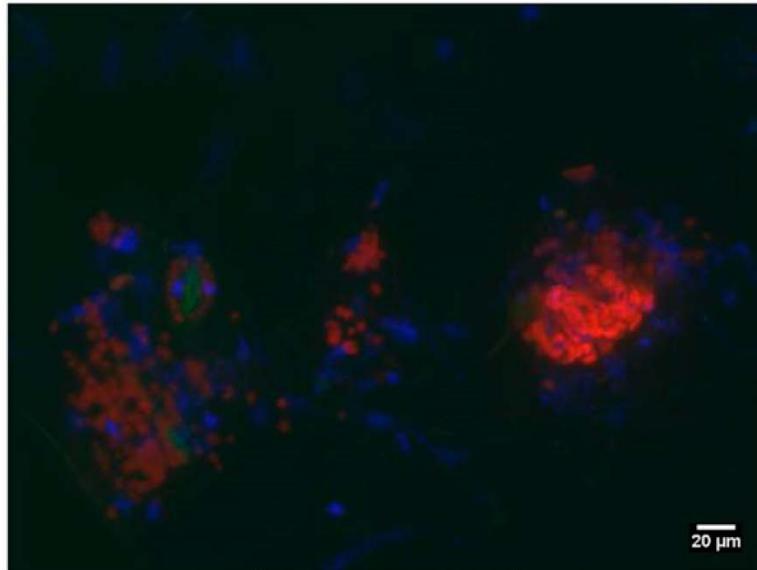
Journal of
Botany

American Journal of Botany: e13–e15, 2011.

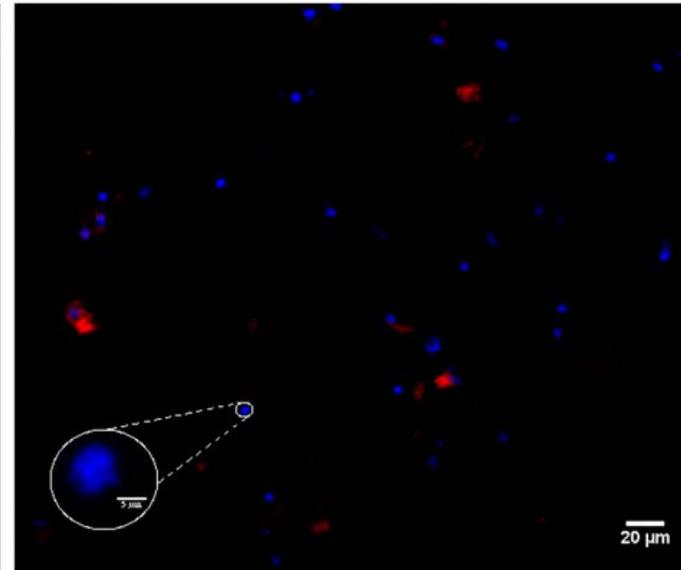
AJB PRIMER NOTES & PROTOCOLS IN THE PLANT SCIENCES

**AN EFFICIENT AND RAPID PROTOCOL FOR PLANT NUCLEAR
DNA PREPARATION SUITABLE FOR NEXT GENERATION
SEQUENCING METHODS¹**

GREGORY CARRIER^{2,6}, SYLVAIN SANTONI³, MARGUERITE RODIER-GOUD¹,
AURÉLIE CANAGUIER⁵, ALEXANDRE DE KOCHKO³, CHRISTINE DUBREUIL-TRANCHANT³,
PATRICE THIS³, JEAN-MICHEL BOURSQUOT³, AND LOÏC LE CUNFF²



total DNA, Vitis



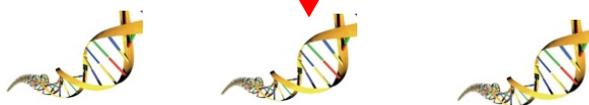
nuclear DNA, Vitis

Stratégie générale

Différents clones



Extraction d'ADN pour NGS



Séquencage en NGS



Stratégie générale

Différents clones



Extraction d'ADN pour NGS



Séquencage en NGS



2009

454

SEQUENCING

Titanium

1,2 M reads

Taille des reads: 400 bases

→ **3 clones**

Stratégie générale

Différents clones



Extraction d'ADN pour NGS



Séquencage en NGS



2009 454

SEQUENCING

Titanium

1,2 M reads

Taille des reads: 400 bases

→ **3 clones**

2010

Solexa™

HiSeq 2000

200 M reads en paired-end

Taille des reads: 100 bases

→ **12 clones**

Stratégie générale

Différents clones



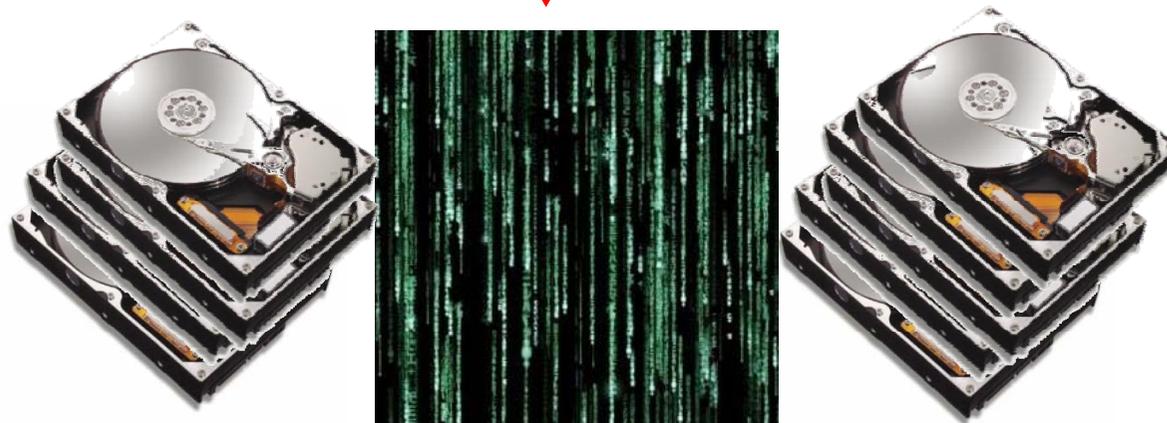
Extraction d'ADN pour NGS



Séquencage en NGS



Analyse des montagnes de données !!!



Stratégie générale

Différents clones



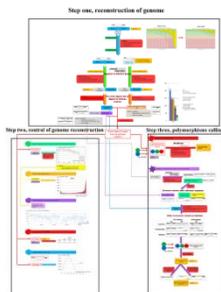
Extraction d'ADN pour NGS



Séquencage en NGS



Analyse des montagnes de données !!!



Pipeline



Serveur



Vincent Maillol

Stratégie générale

Différents clones



Extraction d'ADN pour NGS

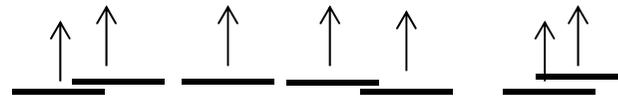


Séquencage en NGS



1-Reconstruire les génomes

Séquence de référence



Alignement sur la référence

Stratégie générale

Différents clones



Extraction d'ADN pour NGS



Séquencage en NGS



1-Reconstruire les génomes

Séquence de référence



Stratégie générale

Différents clones



Extraction d'ADN pour NGS



Séquencage en NGS



Séquence de référence

1-Reconstruire les génomes

2- Comparer les individus



Stratégie générale

Différents clones



Extraction d'ADN pour NGS



Séquencage en NGS



Séquence de référence

1-Reconstruire les génomes

2- Comparer les individus

3-Recherche du polymorphisme

	Clone1 (AA)
SNP. Chr1 Pos. 10	Clone2 (AA)
	Clone3 (AT)

Comparaison des méthodologies de séquençage nouvelle génération

454
SEQUENCING

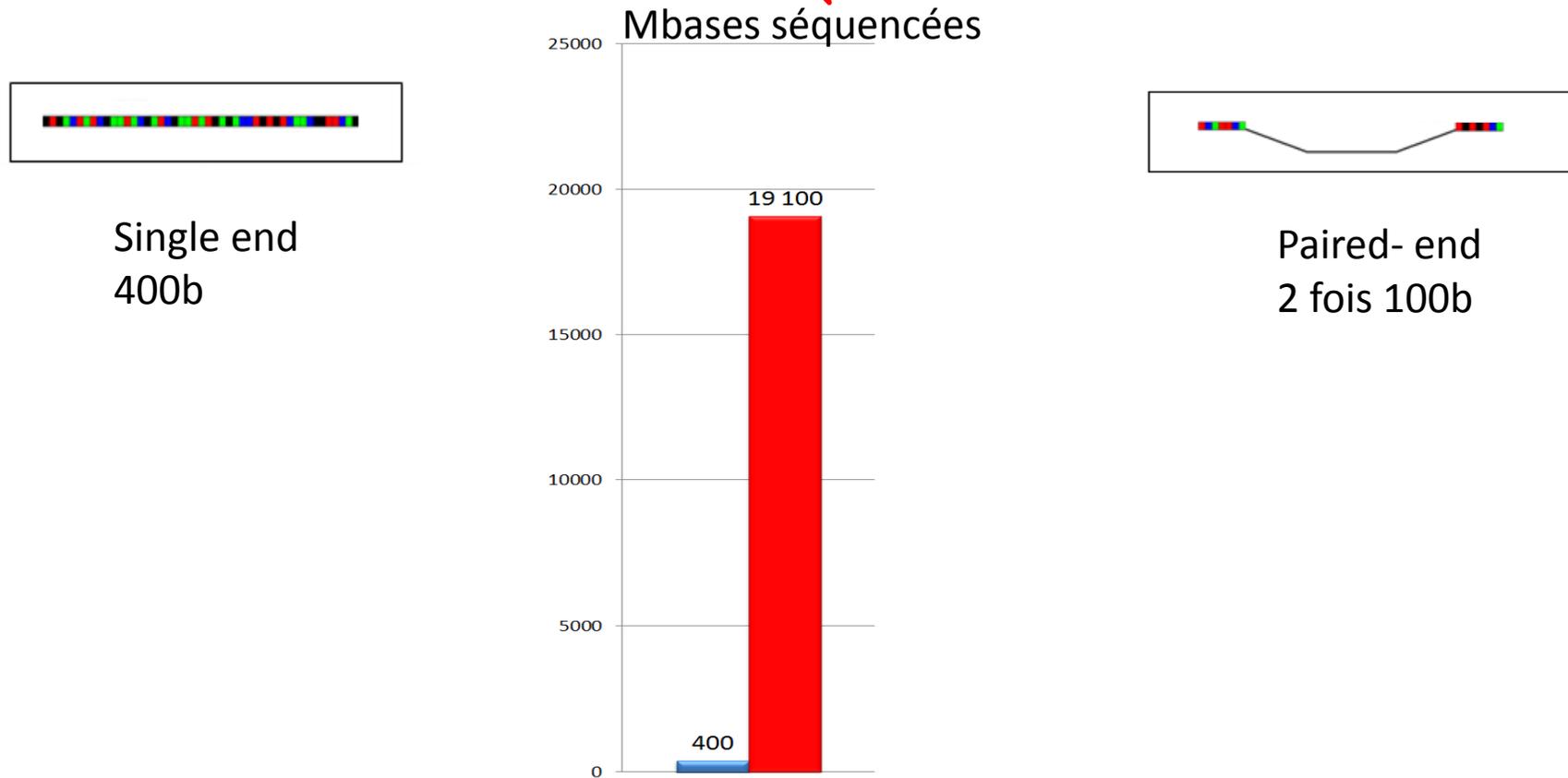


V.S

Solexa™



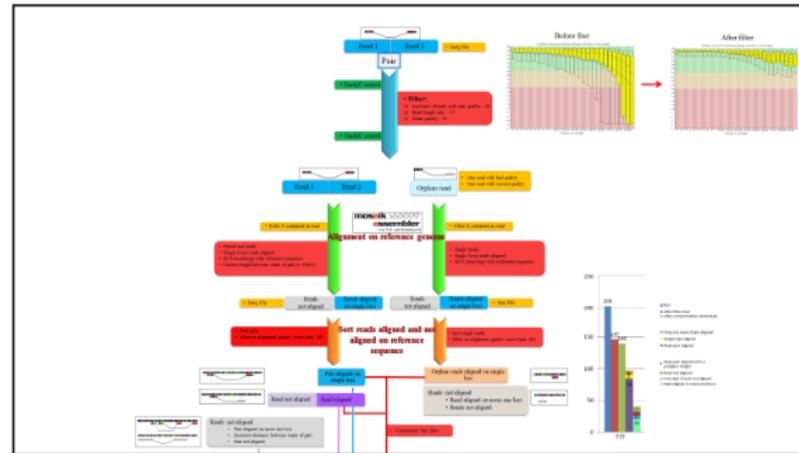
Données produites



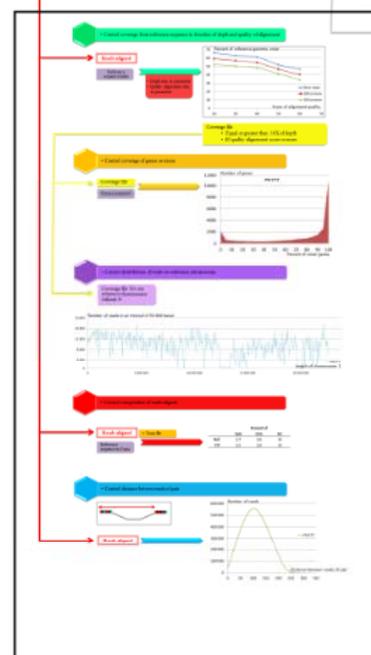
Solexa, 50 fois plus productif que le 454

Analyse des séquences : Bacchus pipeline

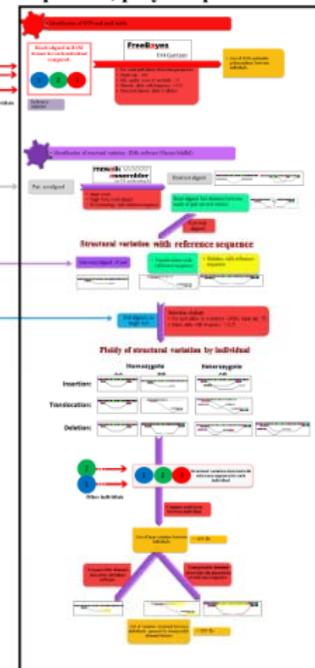
Step one, reconstruction of genome



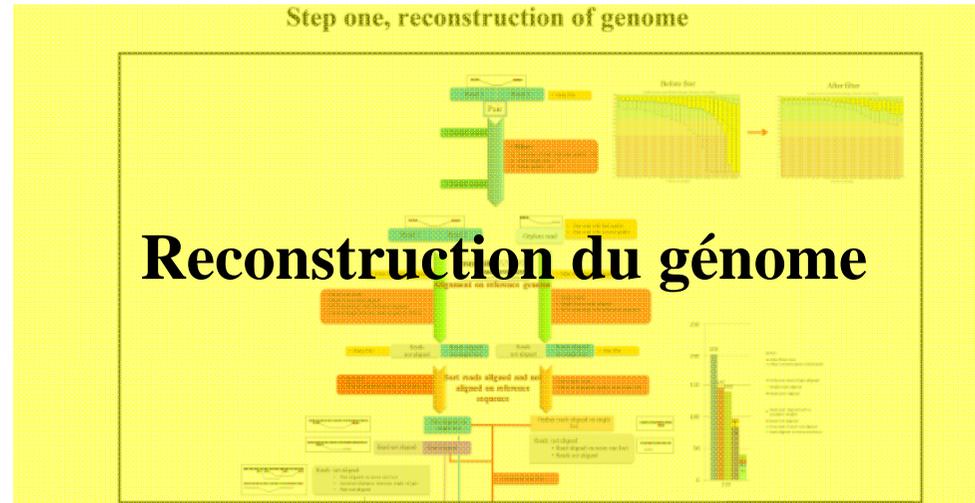
Step two, control of genome reconstruction



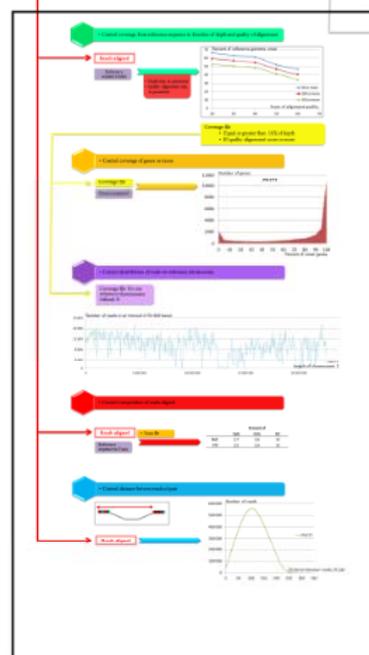
Step three, polymorphisms calling



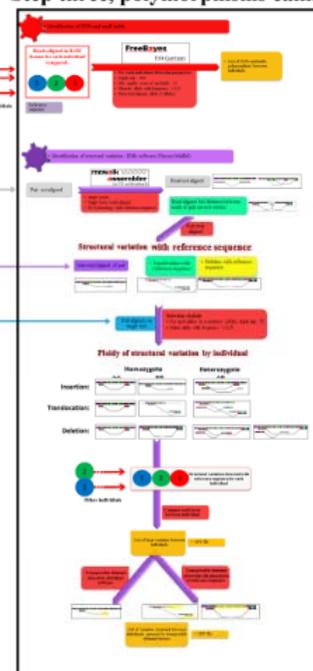
Bacchus pipeline



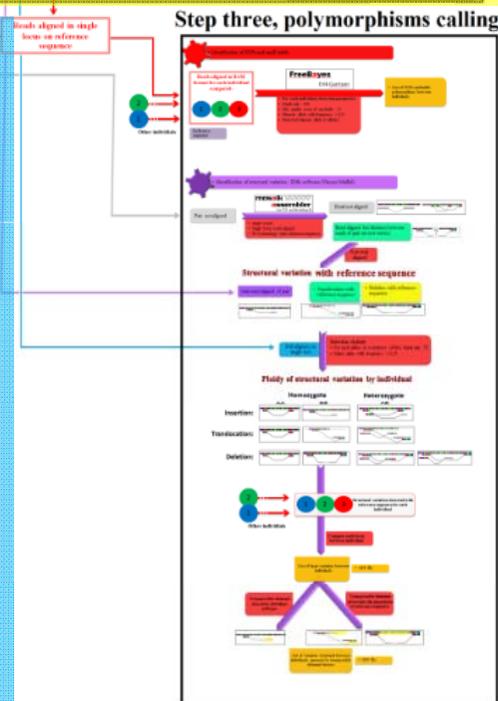
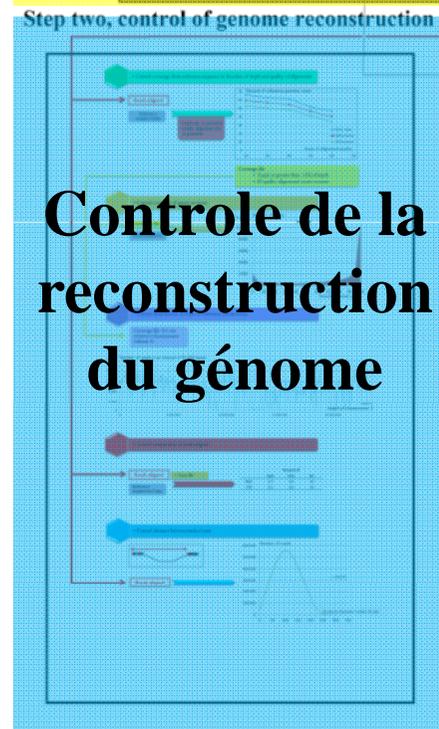
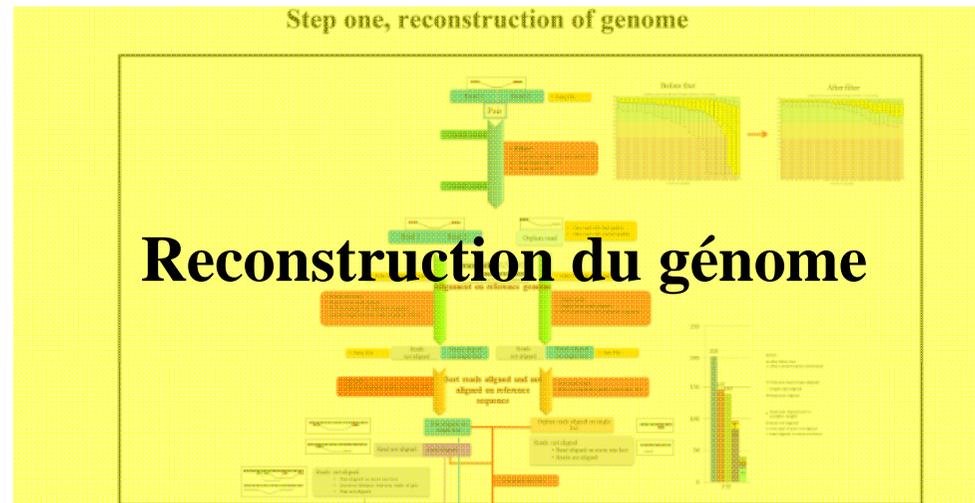
Step two, control of genome reconstruction



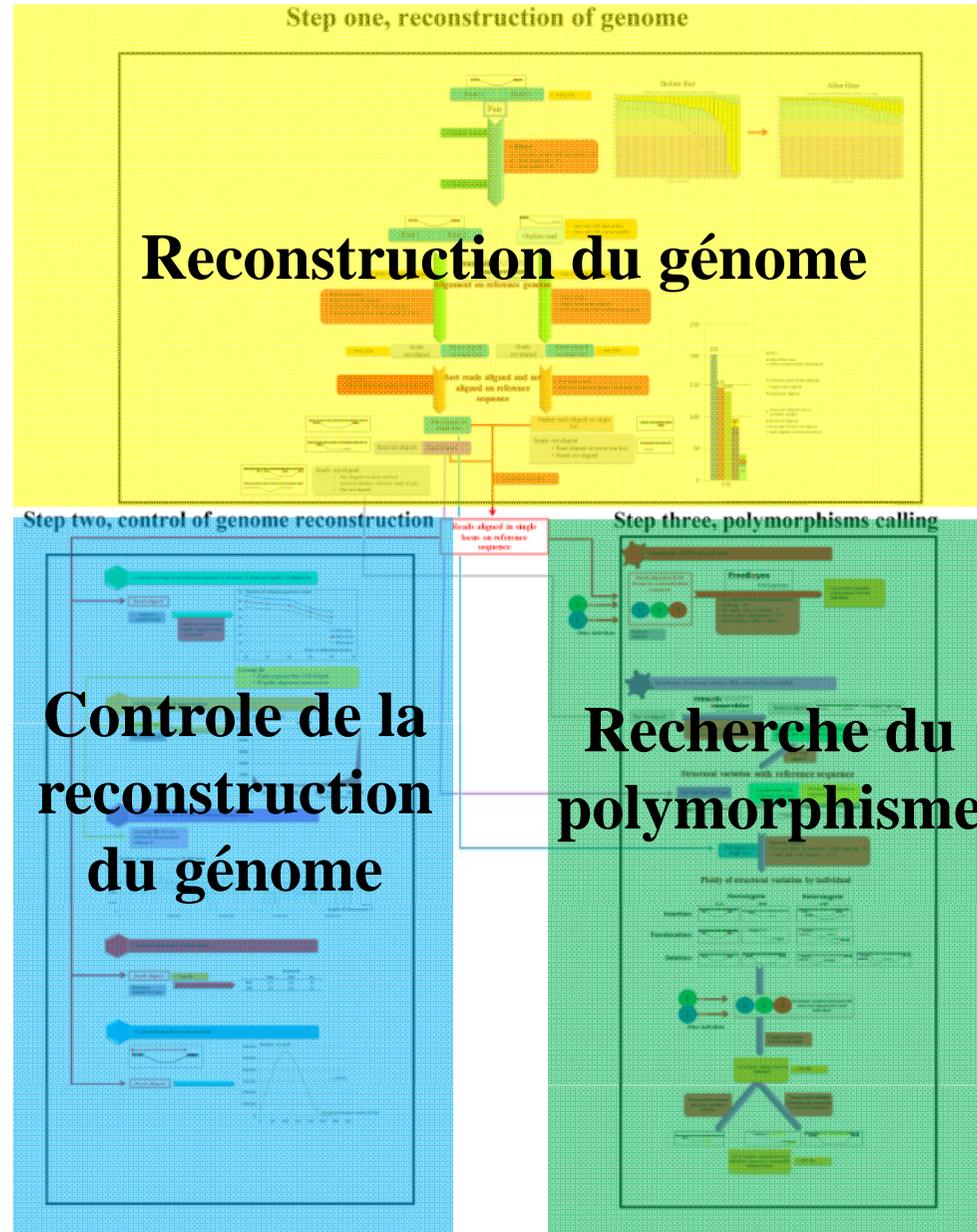
Step three, polymorphisms calling



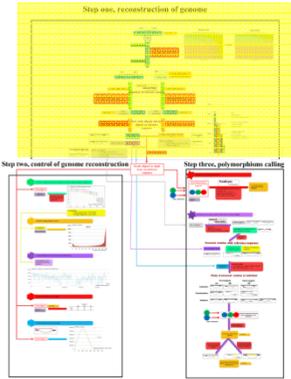
Bacchus pipeline



Bacchus pipeline



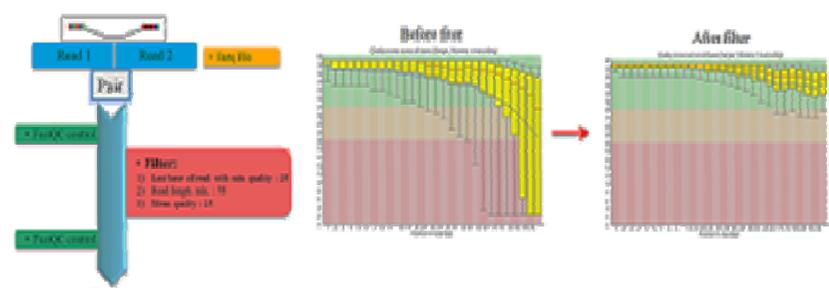
Reconstruction du génome



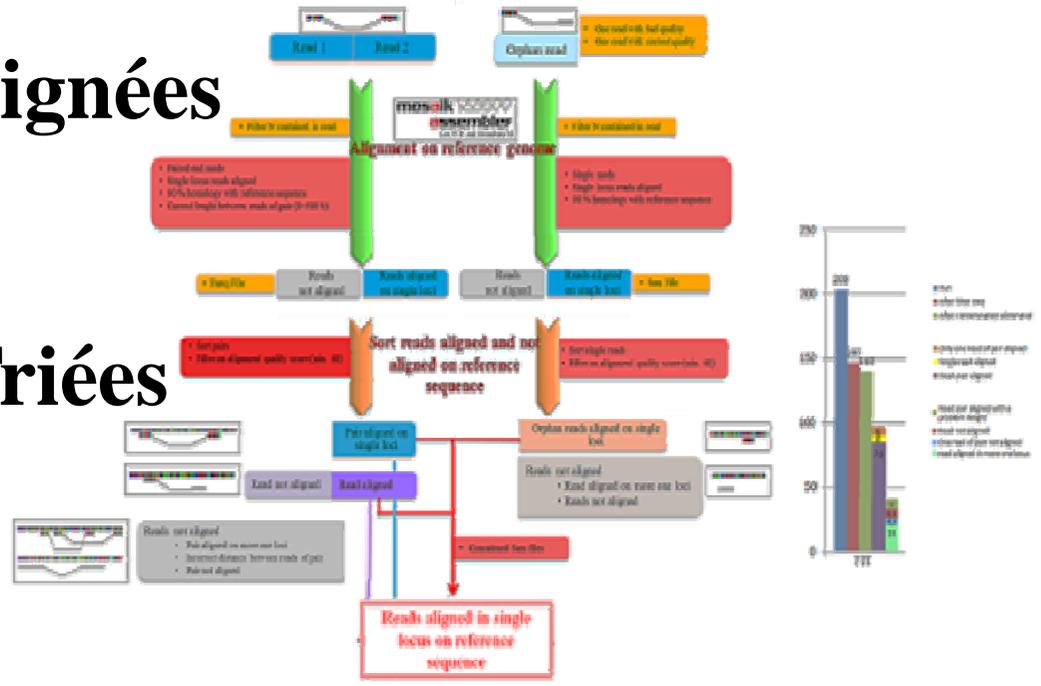
Séquences



1) Filtrées

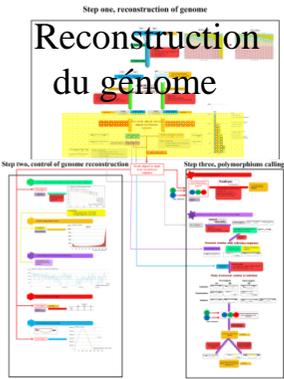


2) Alignées

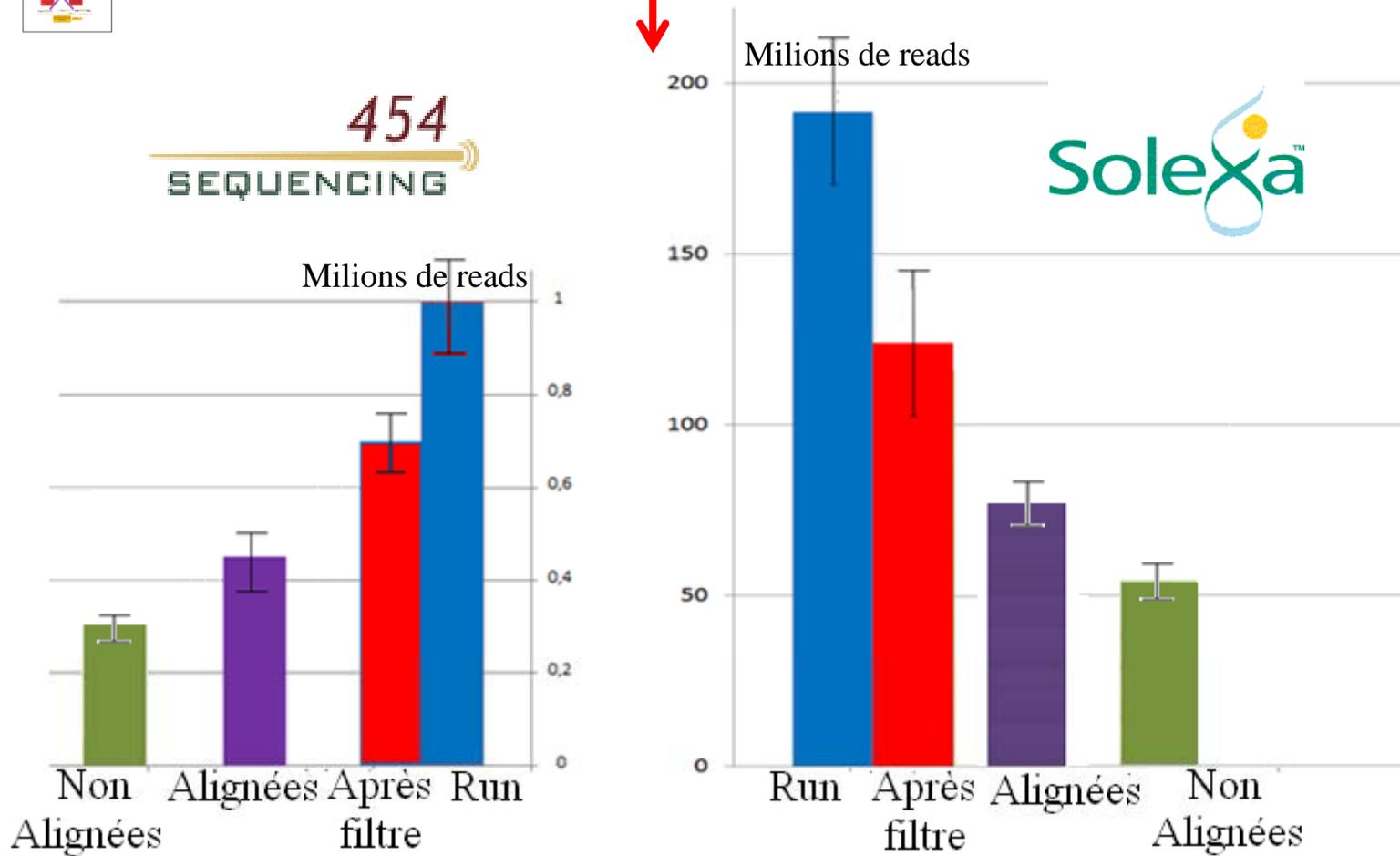


3) Triées

Reconstruction du génome

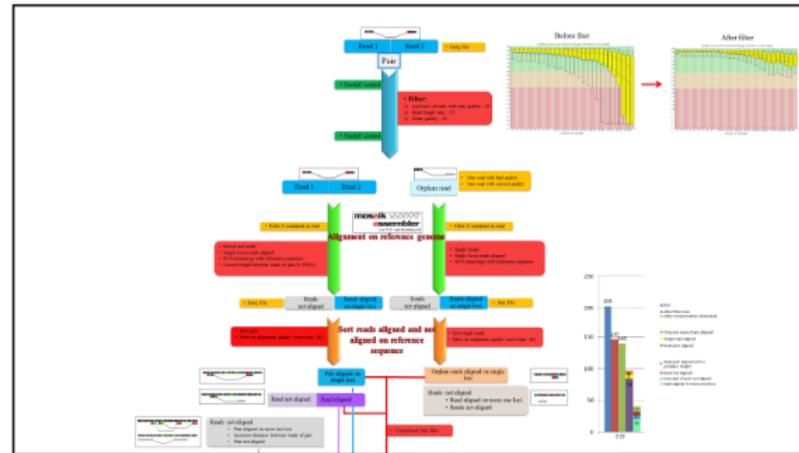


3) Triées

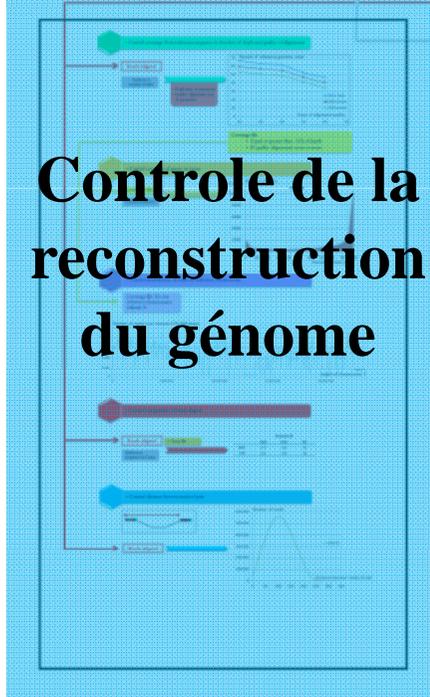


Bacchus pipeline

Step one, reconstruction of genome

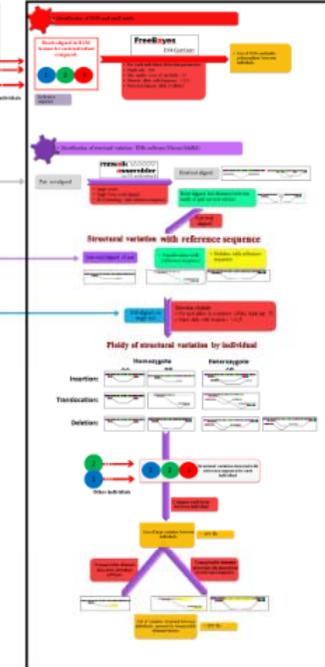


Step two, control of genome reconstruction

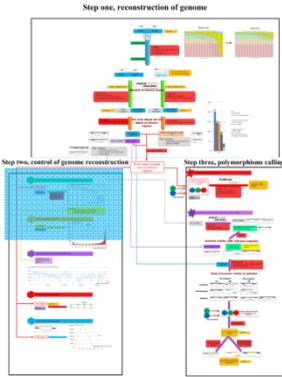


Controle de la reconstruction du génome

Step three, polymorphisms calling

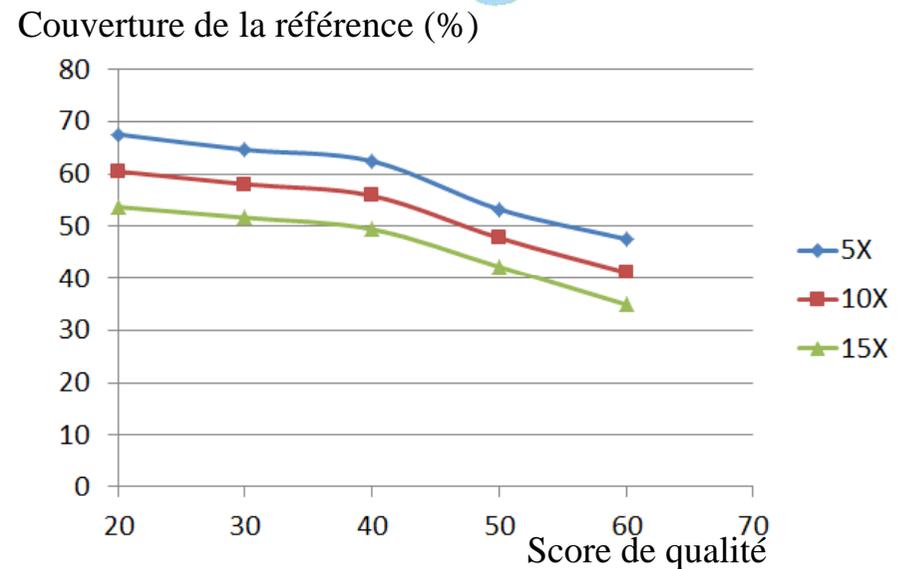
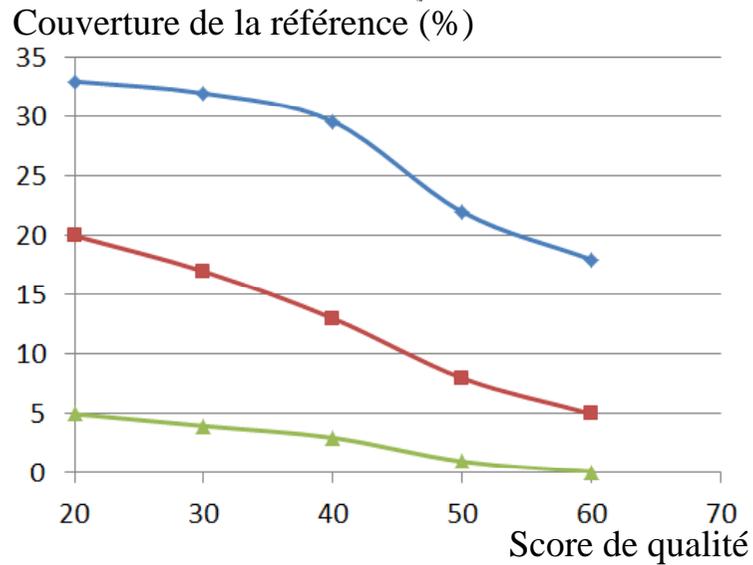


Contrôle de la reconstruction du génome

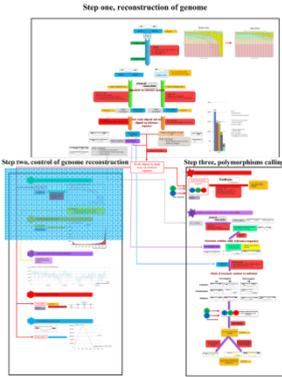


Couverture du génome en fonction de - la qualité d'alignement
- la profondeur

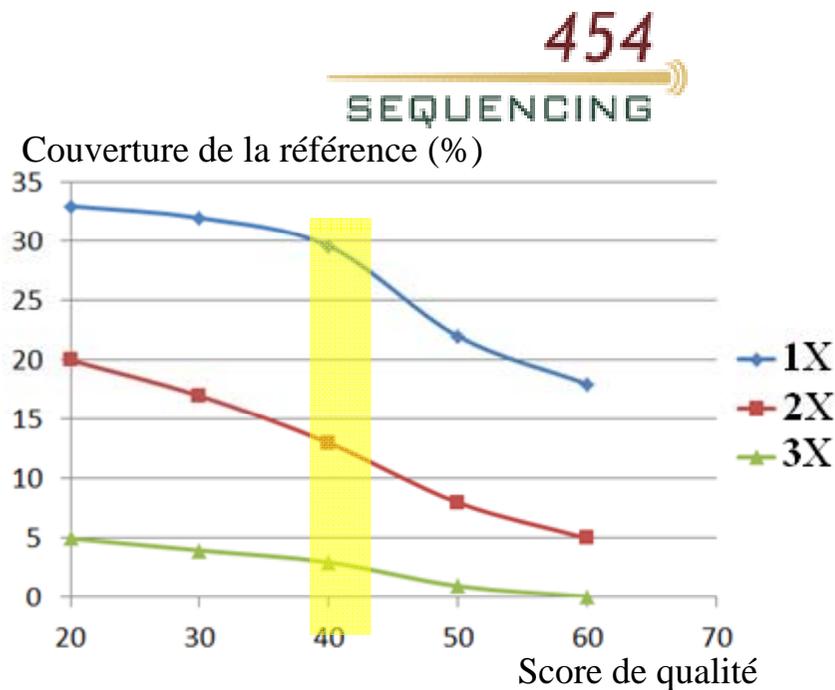
454
SEQUENCING



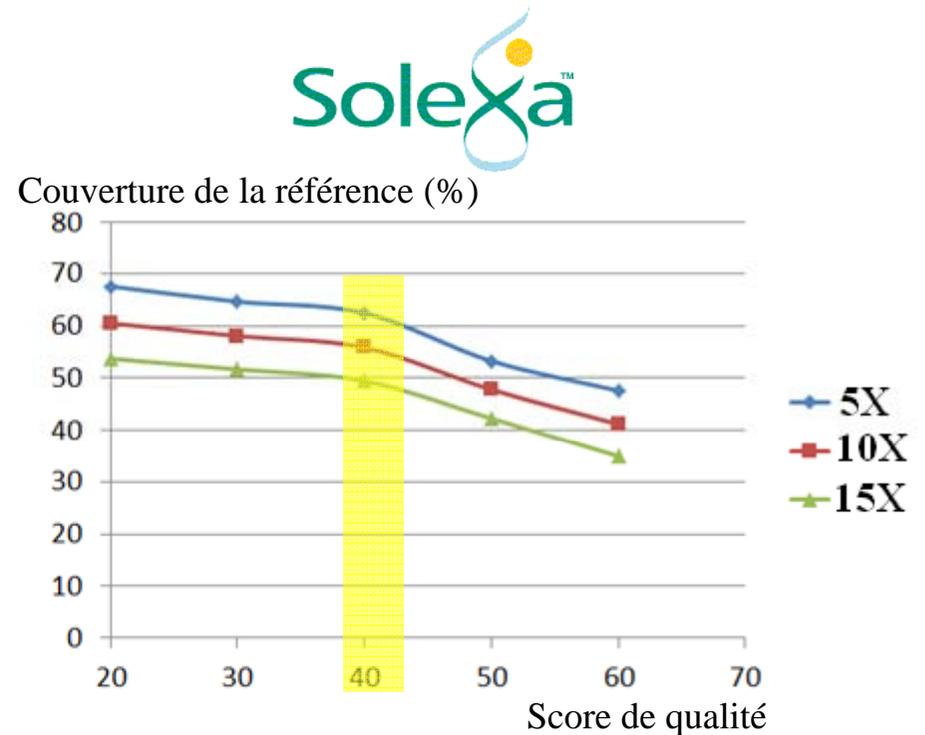
Contrôle de la reconstruction du génome



Couverture du génome en fonction de - la qualité d'alignement
- la profondeur

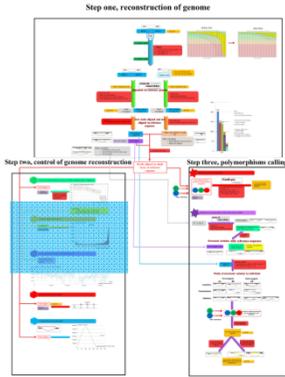


1% du génome comparé en 6X



50% du génome comparé en 10X

Contrôle de la reconstruction du génome

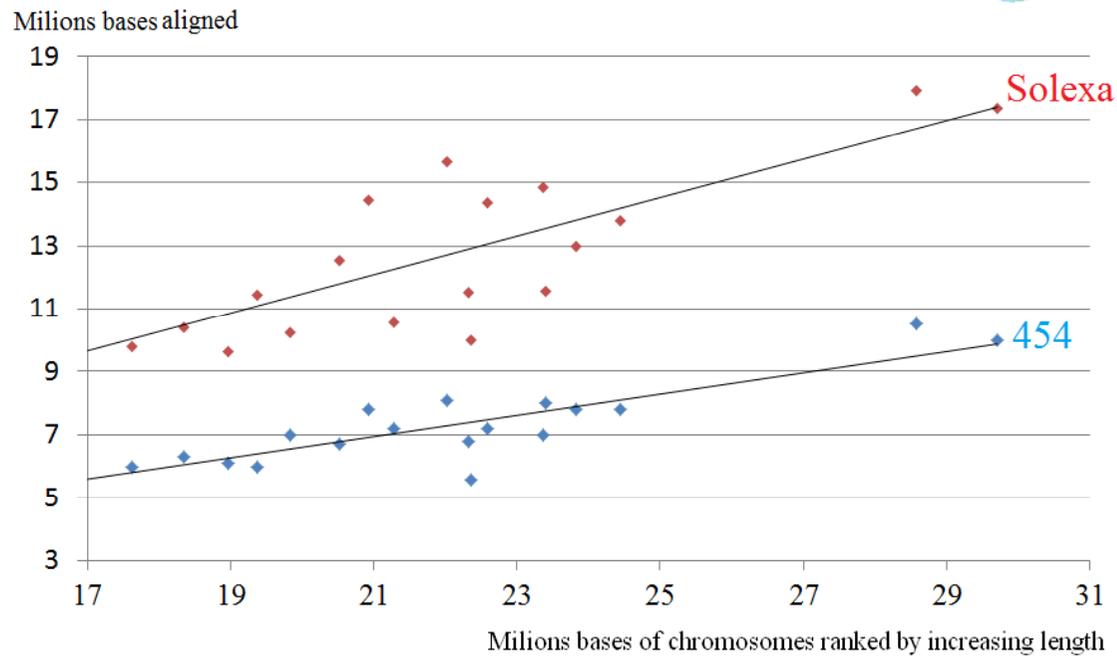


Distribution des reads dans le génome

454
SEQUENCING

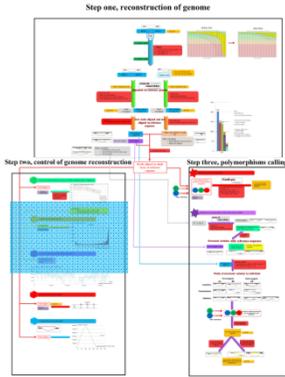


$R^2 > 0,80$



Distribution des séquences sans *a priori* sur le génome

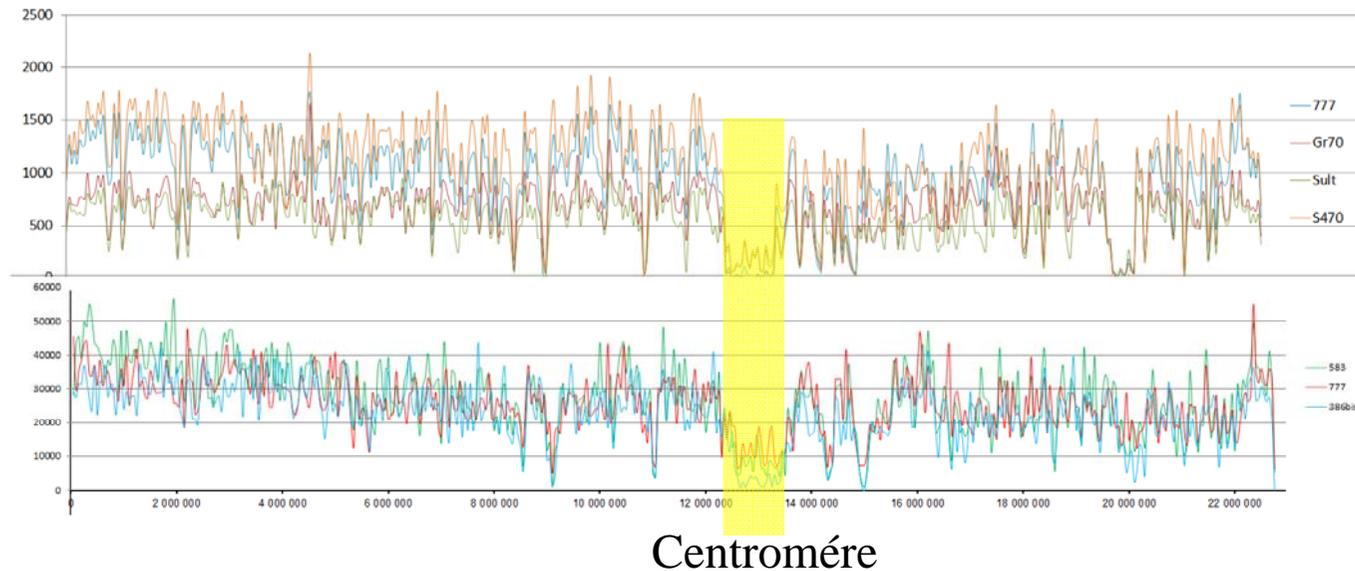
Contrôle de la reconstruction du génome



Distribution des reads sur les chromosomes

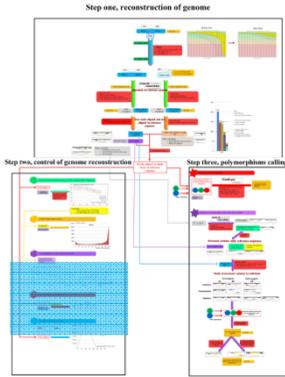
Solexa™

454
SEQUENCING

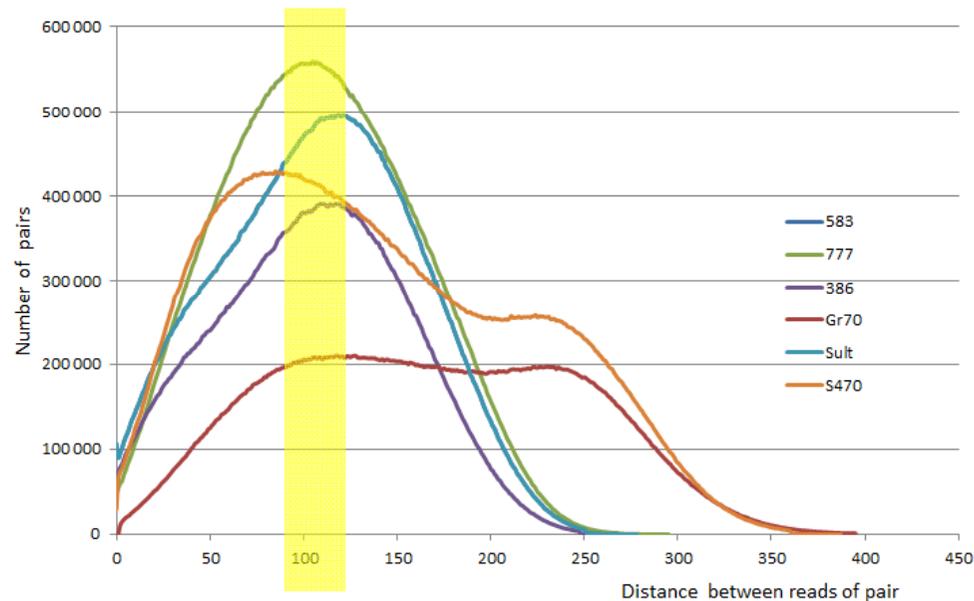


Difficile d'accéder aux séquences répétées (Pvalue < 0,05)

Contrôle de la reconstruction du génome

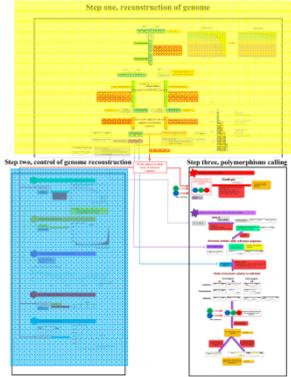


Distance des reads constituant les paires



Chevauchement des reads de la paire

Contrôle de la reconstruction du génome



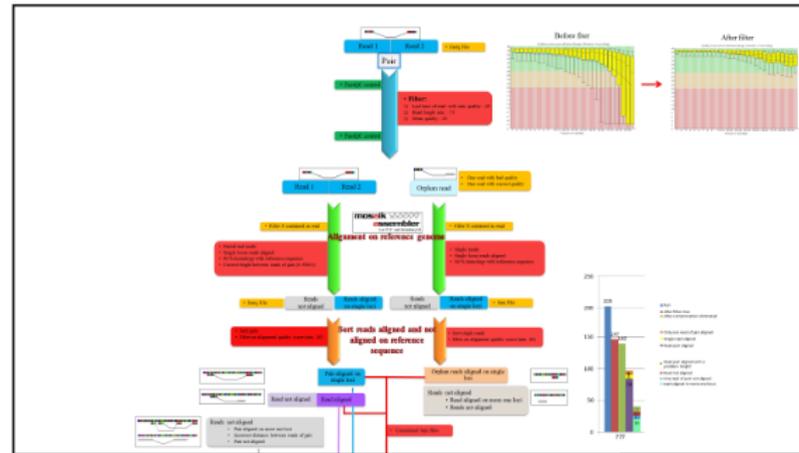
Bilan de la reconstruction du génome



Pas de différence majeure...
Mais Solexa 50 fois plus productif

Bacchus pipeline

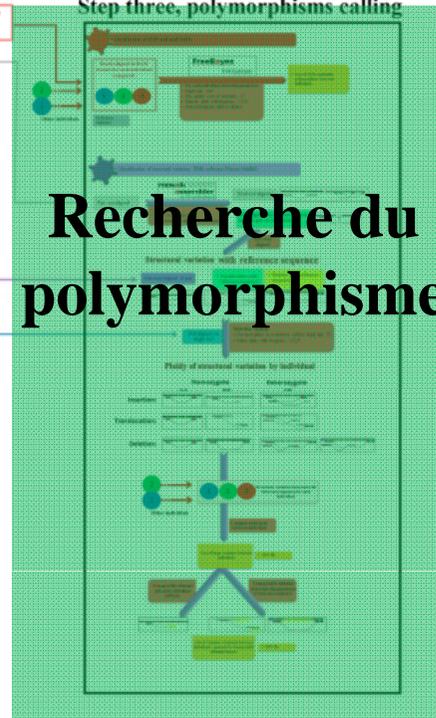
Step one, reconstruction of genome



Step two, control of genome reconstruction

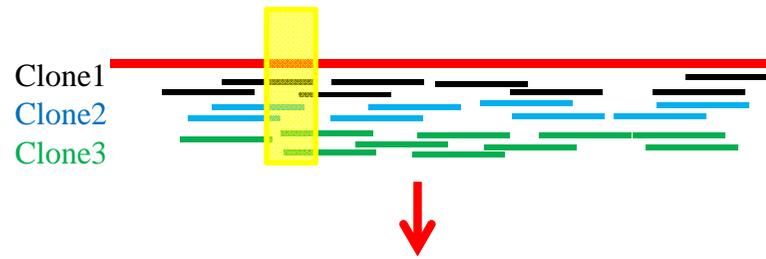
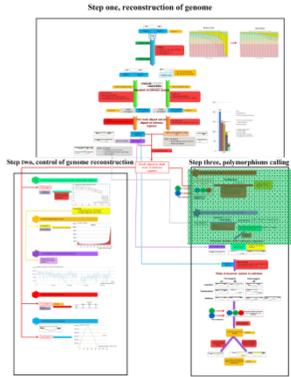


Step three, polymorphisms calling



Recherche de polymorphisme

Type SNPs et Indels



FreeBayes
Erik Garrison

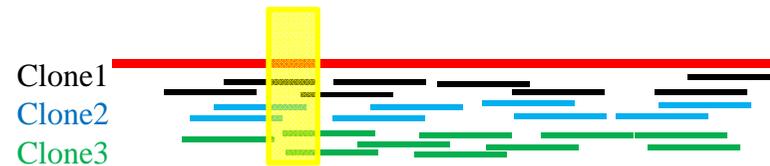
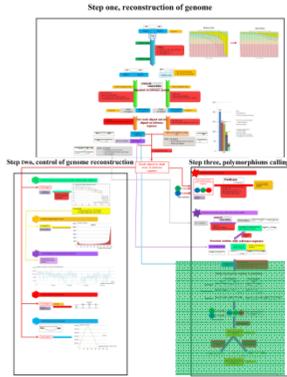
454
SEQUENCING

Solexa™

- SNPs	OK	OK
- Indels (-20 b)	OK	OK
- Indels (20 - 100 b)	OK	(OK)
- SSRs (100- 250 b)	OK	Non
- 3 allèles	Non	OK
- Faux positif	Problème homopolymère	/

Recherche de polymorphisme

Polymorphisme structural



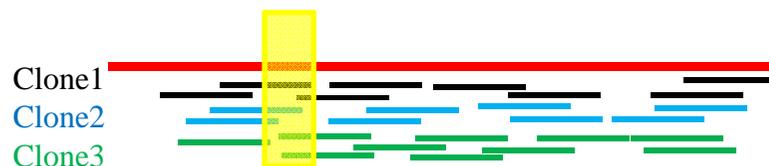
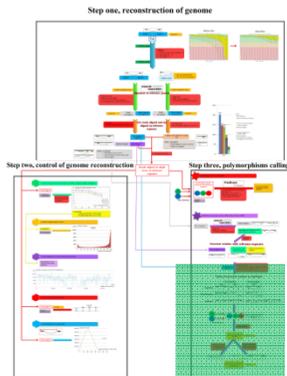
454
SEQUENCING

Solexa™

- Elements tranposables connus	OK	OK
- Délétions / insertions de plus de 1Kb	non	OK
- Translocations	non	OK
- Inversions	non	OK

Recherche de polymorphisme

Polymorphismes structural



454
SEQUENCING

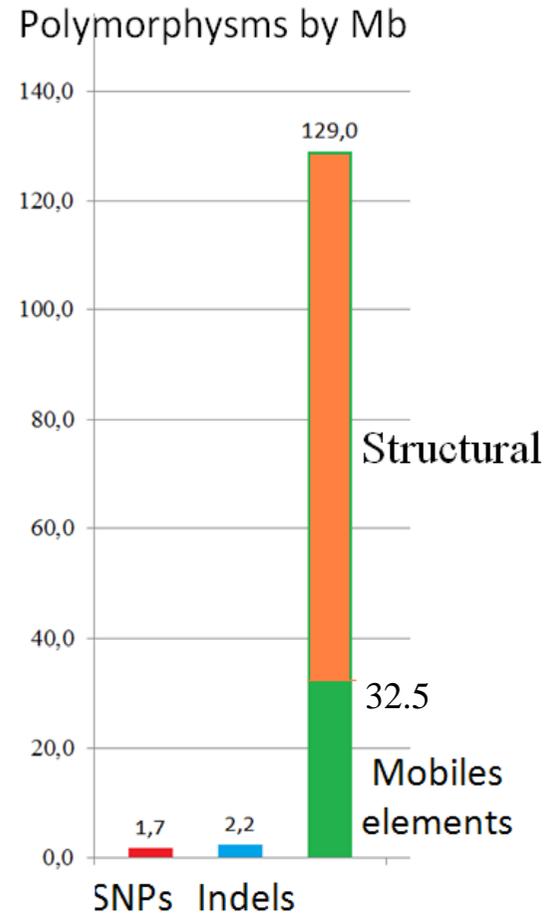
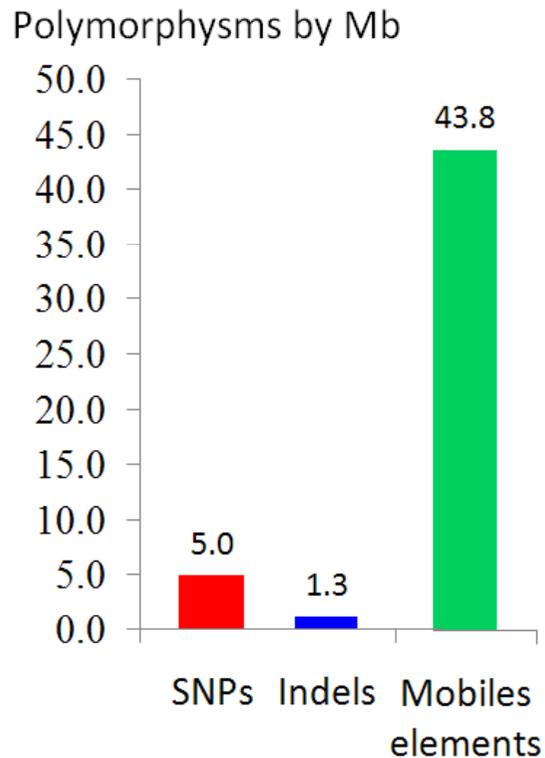


- Elements tranposables connus	OK	OK
- Délétions / insertions de plus de 1Kb	non	OK
- Translocations	non	OK
- Inversions	non	OK

(Mais si 454 en paired end → OK)

Résultats polymorphismes entre clones

454
SEQUENCING



Assez proche- complémentaire

Bilan de la comparaison



- **Quantité de données** Faibles (1% - 6X) Fortes (50% - 20X)
 Prob. Homopolymère
- **Qualité des données** Equivalente
- **Polymorphisme** Complémentaire

Bilan de la comparaison



- **Quantité de données** Faibles (1% - 6X) Fortes (50% - 20X)
 Prob. Homopolymère
- **Qualité des données** Equivalente
- **Polymorphisme** Complémentaire

Le séquenceur 3ème génération idéal ?

→ fort débit, séquences longues

MERCI à VOUS ET ..

Loïc Le Cunff

Patrice This et Jean-Michel Boursiquot

Laurent Audeguin

Vincent Maillol

Maud Pajeile

L'UMT GénoVigne

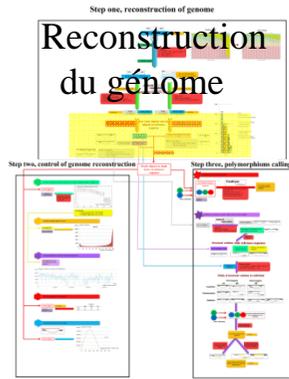
L'équipe DAVEM-vigne

L'équipe de l'IFV

La plateforme de Génotypage de AGAP
Génotoul de Toulouse

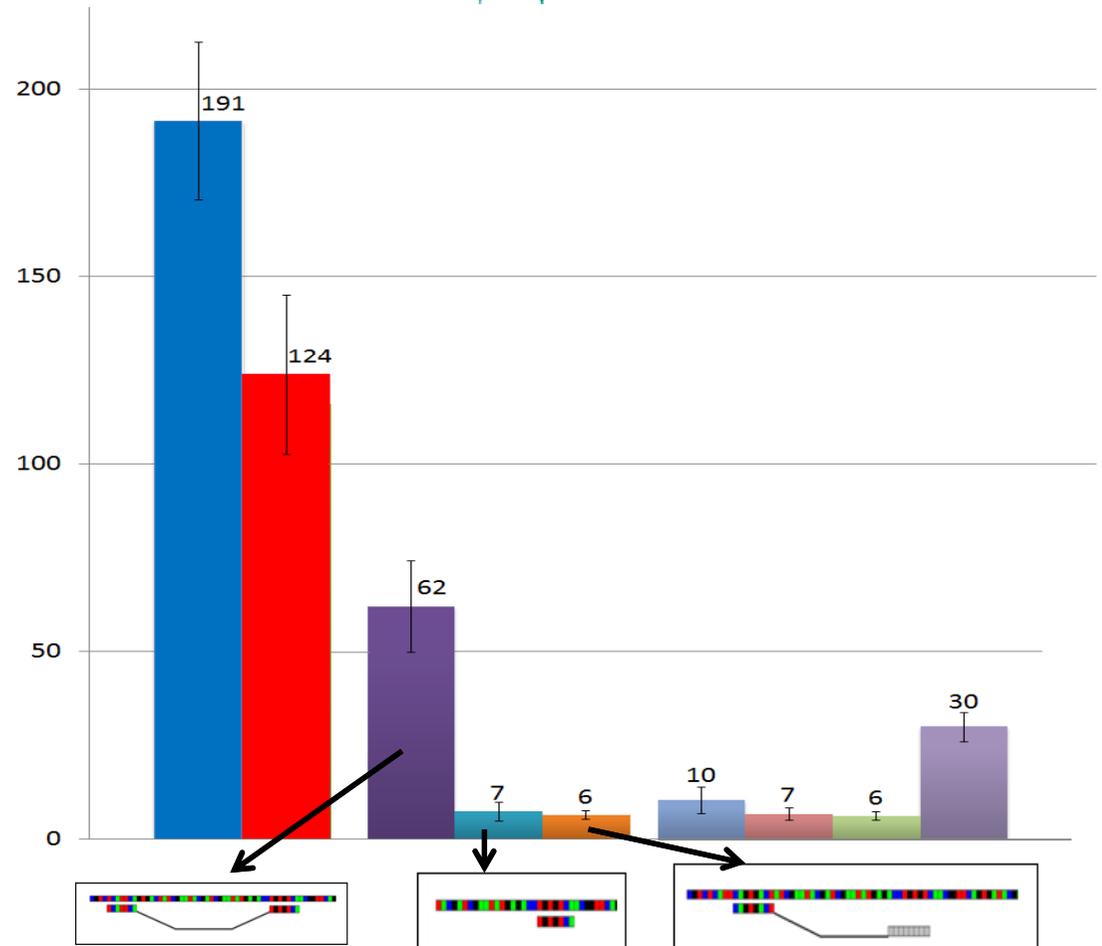
Les conservatoires de Clones

Bacchus pipeline

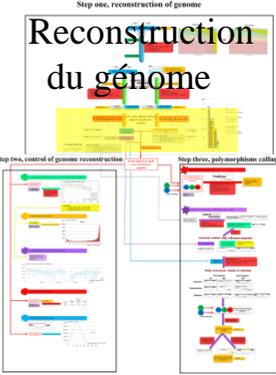


3- Tri des séquences

454
SEQUENCING



Bacchus pipeline



3- Tri des séquences

454
SEQUENCING

