

EPGV Meeting
12 mai 2009

454 / Analyses bioinformatiques

François Artiguenave

Méthode de séquençage: Pyrosequencing.

Amplification: Emulsion PCR sur billes d'agarose.

Taille de la lecture: 250-550 bp

Temps pour 1 run: ~ 8h

Nombre de lecture : 500 000

Mb/run: >100Mb

Couverture optimale : Entre x10 et x15

Avantages:

- Incorporation pas à pas entraîne peu d'erreurs de substitution
- Pas de biais de clonage

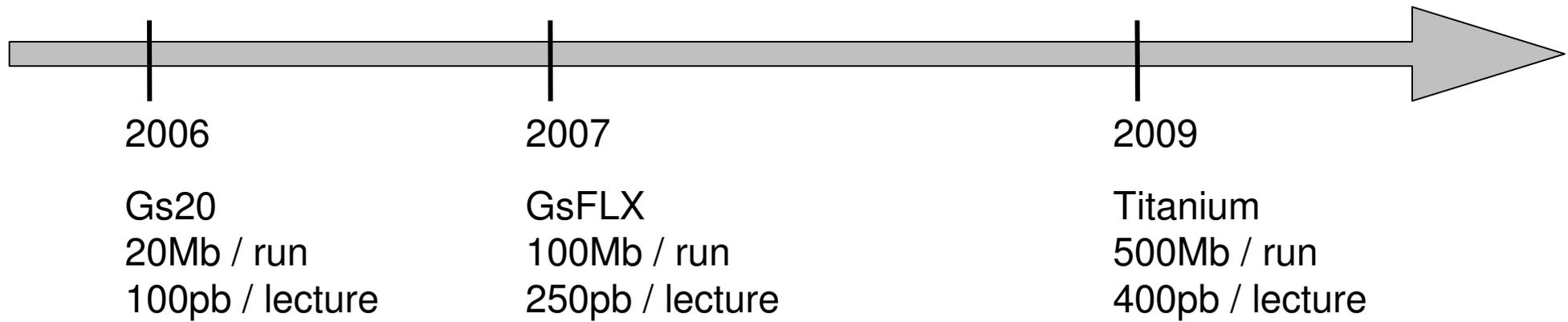
Inconvénients:

- Coût de l'expérimentation par rapport à l'utilisation du Solexa et SOLiD
- Erreurs au niveau des régions contenant des homopolymères





454 / Roche – Genome Sequence FLX

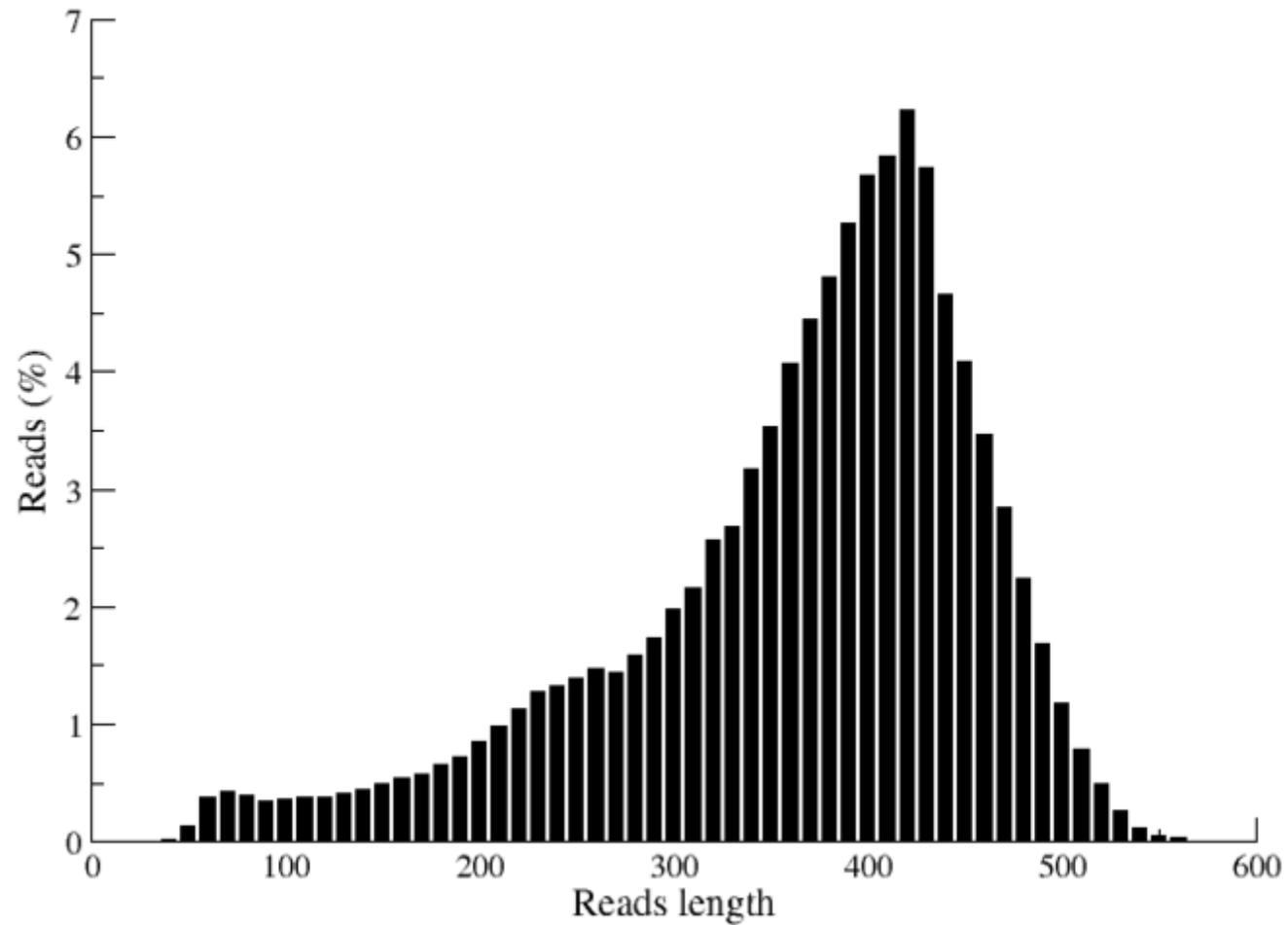


Séquençage de génomes



454 / Roche – Genome Sequence TITANIUM

✓ Distribution des tailles de lecture

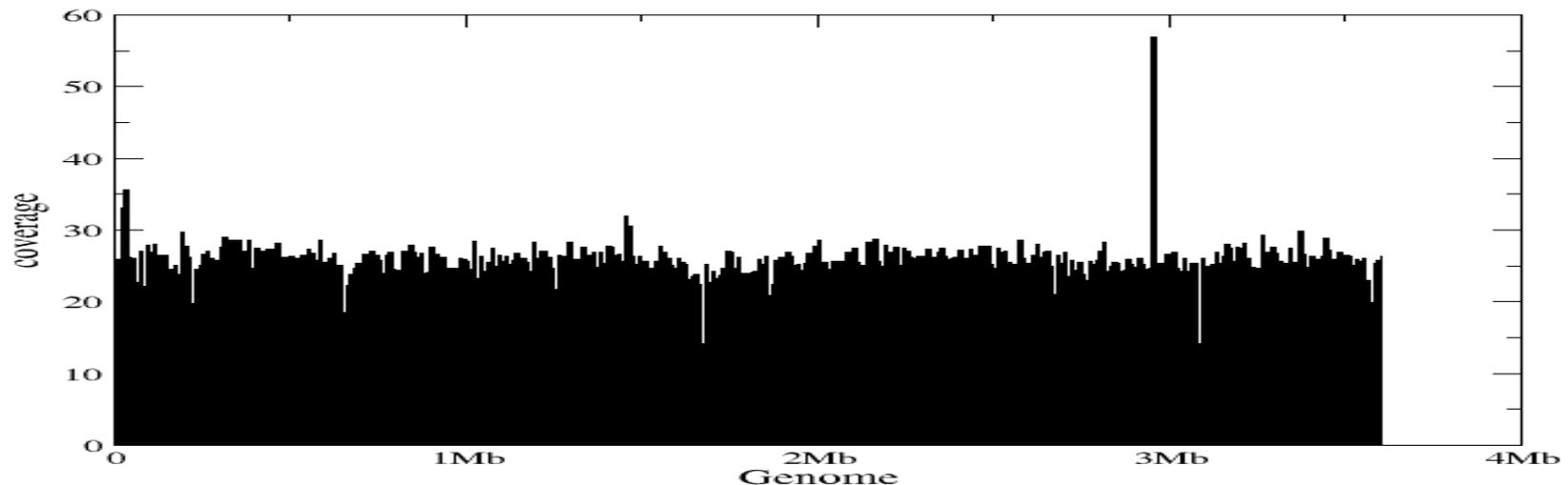


Séquençage de génomes



454 / Roche – Genome Sequence FLX

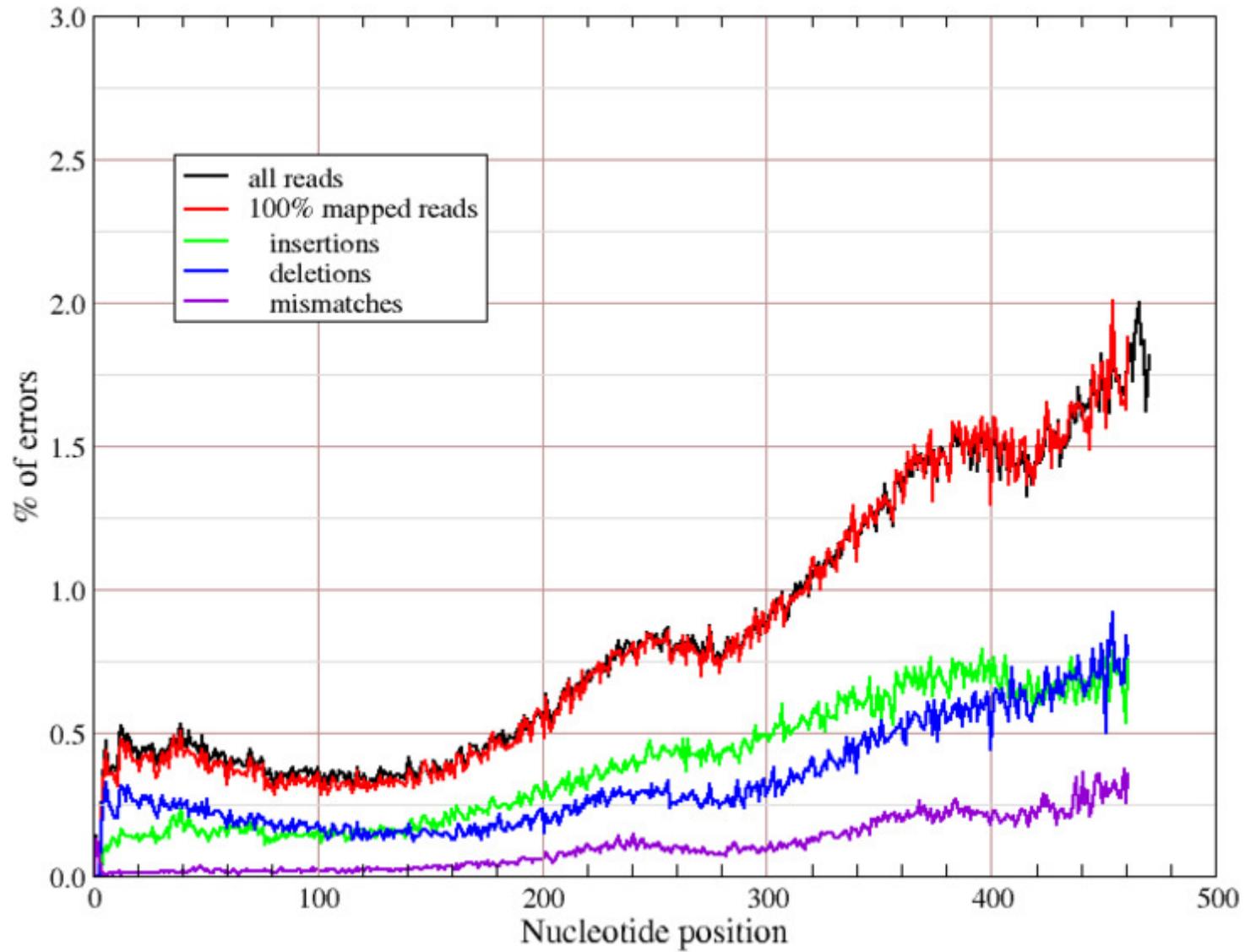
- ✓ Alignement des lectures au niveau nucléotidique : 521.193 lectures mappées (soit 99,68%)
- ✓ 93.553.967 nt alignés contenant 800.295 erreurs (soit $8,6 \cdot 10^{-3}$ erreurs, de l'ordre de 10^{-3} à $5 \cdot 10^{-3}$ en sanger)
- ✓ 17% délétions, 62% insertions, 21% mismatches (12% de Ns).
- ✓ Erreurs concentrées autour des régions homopolymériques => le taux d'erreur n'est pas constant, il dépend du taux d'homopolymères



Séquençage de génomes



454 / Roche – Genome Sequence FLX

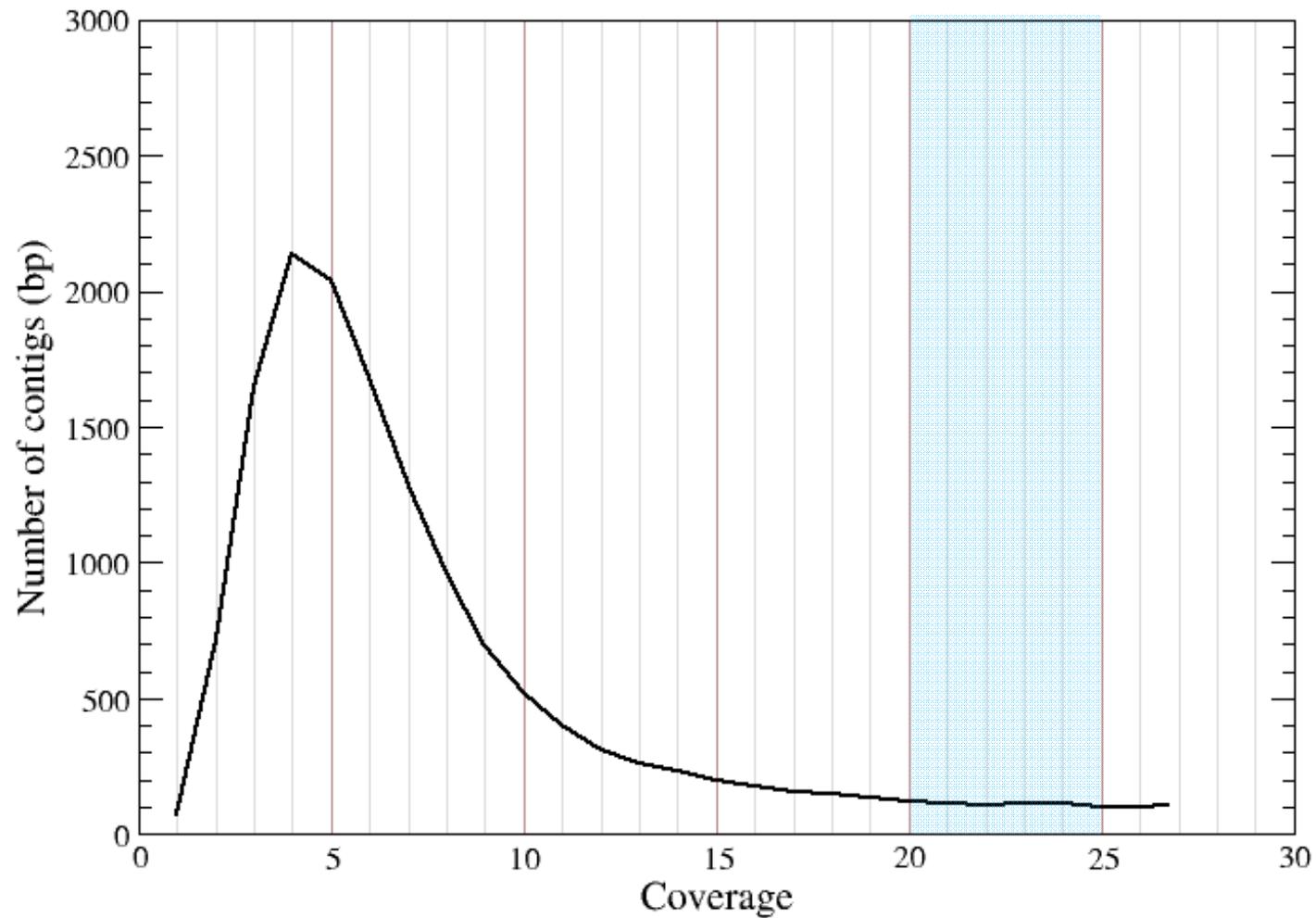


Séquençage de génomes



454 / Roche – Genome Sequence FLX

✓ Quelle profondeur de séquençage nécessaire ?

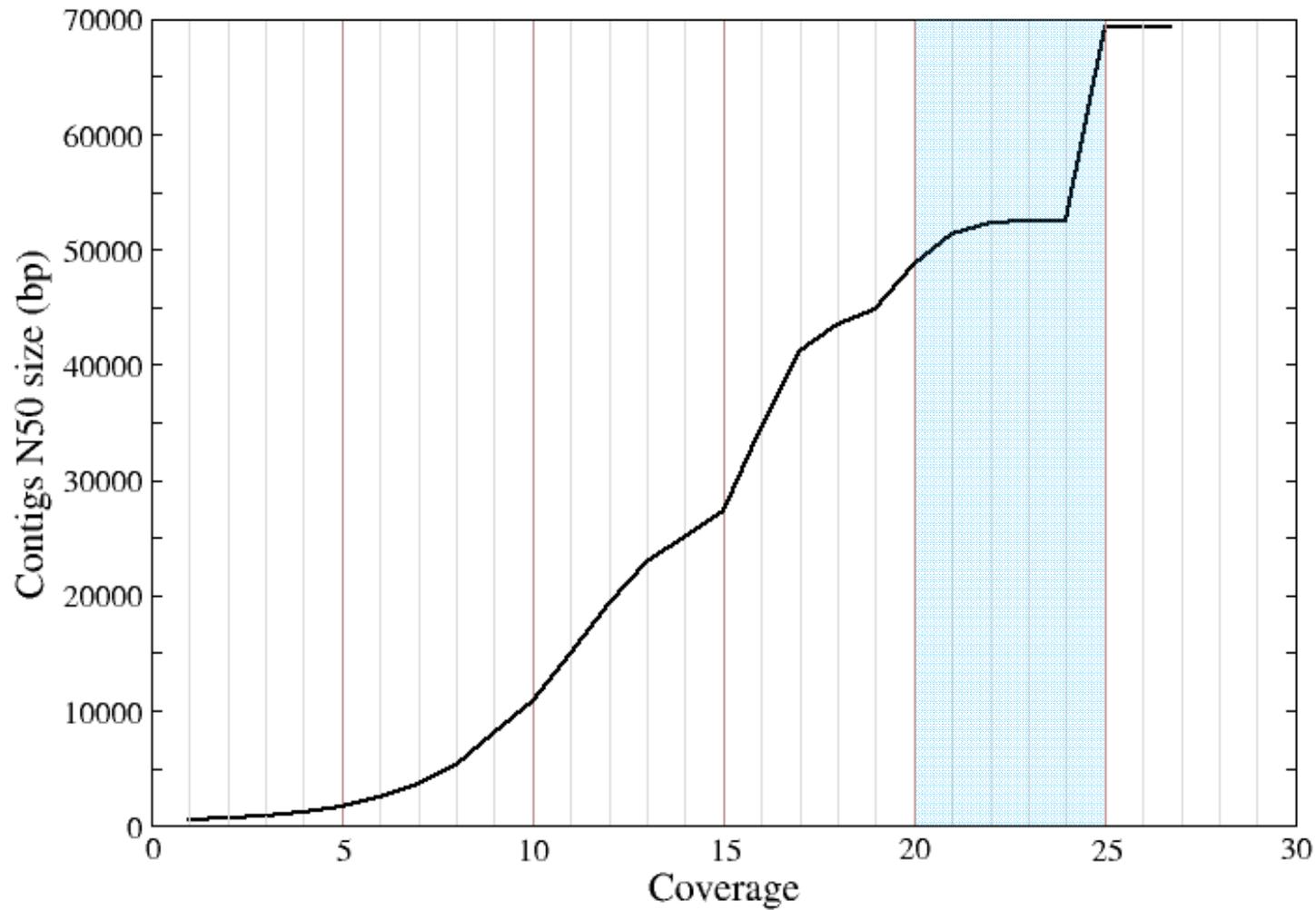


Séquençage de génomes procaryotes



454 / Roche – Genome Sequence FLX

✓ Quelle profondeur de séquençage nécessaire ?

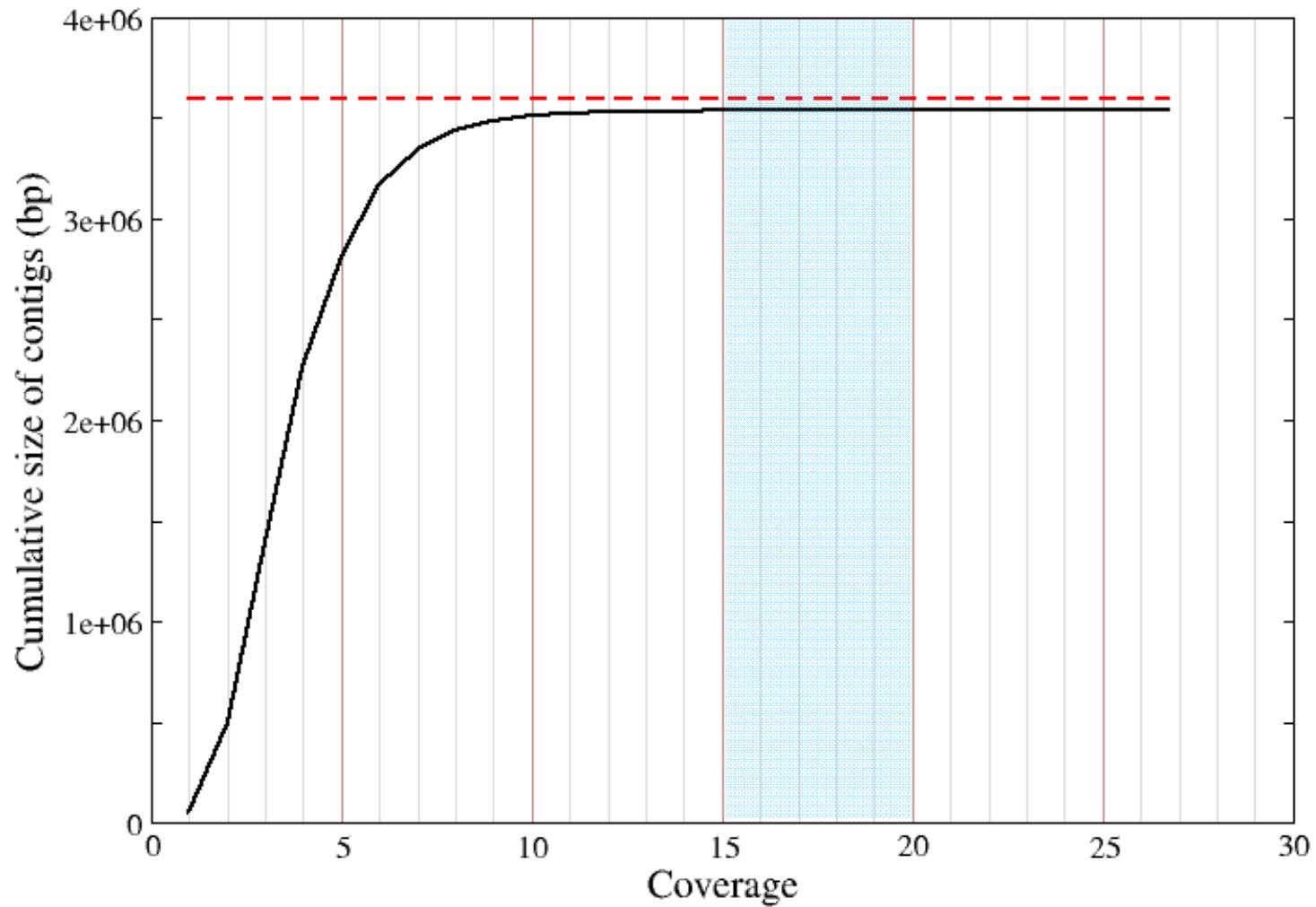


Séquençage de génomes



454 / Roche – Genome Sequence FLX

✓ Quelle profondeur de séquençage nécessaire ?



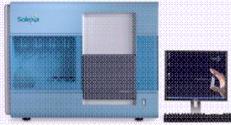
Séquençage de génomes

	Sanger	Unpaired 454	Unpaired + PE 454
Coverage	7.4X	20X	25X
Assembler	Arachne (Broad Institute)†	Newbler (454/Roche)†	Newbler (454/Roche)†
# of contigs	173	119	119
Contigs N50 (Kb)†	39.0	48.7	58.2
# of scaffolds	2	119	10
Scaffolds N50 (Kb)†	2,200	48.7	1,000
Assembly size (% of reference)†	3.417Mb (95%)†	3.542 Mb (98%)†	3.544 Mb (98%)†
Mis-assemblies	0	0	0
# of errors	3,442	420	431
Substitutions	2,494	67	75
Insertions / Deletions	948	353	356

Séquençage de génomes

- Structure de l'assemblage satisfaisante (plus de scaffolds => banque de 3 et 10Kb pour le sanger contre 3Kb pour 454 PE)
- Meilleure représentation du génome (couverture homogène)
- Taux d'erreurs : ~ 1 erreur / 8,5Kb,
- de nombreux indels (problème pour la prédiction de gènes)
- La qualité peut être améliorée par l'ajout de lectures présentant un type d'erreur différent pour corriger les indels de l'assemblage 454

Séquençage de génomes



Illumina / Solexa – Genetic Analyzer 1G

- ✓ Test sur *Acinetobacter baylyi* (3,5Mb) :
- environ 10M de lectures de 36pb
 - taille cumulée de 440Mb, soit 120 équivalents génome (120X)

	Lane6	Lane7	Lane8
Mapped reads	87,75%	89,74%	88,76%
100% (length) mapped reads	78,38%	81,05%	79,80%
Perfect reads	68,17%	71,41%	69,27%
# of aligned bases	85.357.257	139.282.306	163.453.711
# of errors	321.670 (0,377%) [†]	483.135 (0,347%)[†]	631.224 (0,386%) [†]
Insertions	0,17%	0,15%	0,14%
Deletions	1,02%	1,03%	0,99%
Mismatches	98,81%	98,81%	98,87%

Séquençage de génomes

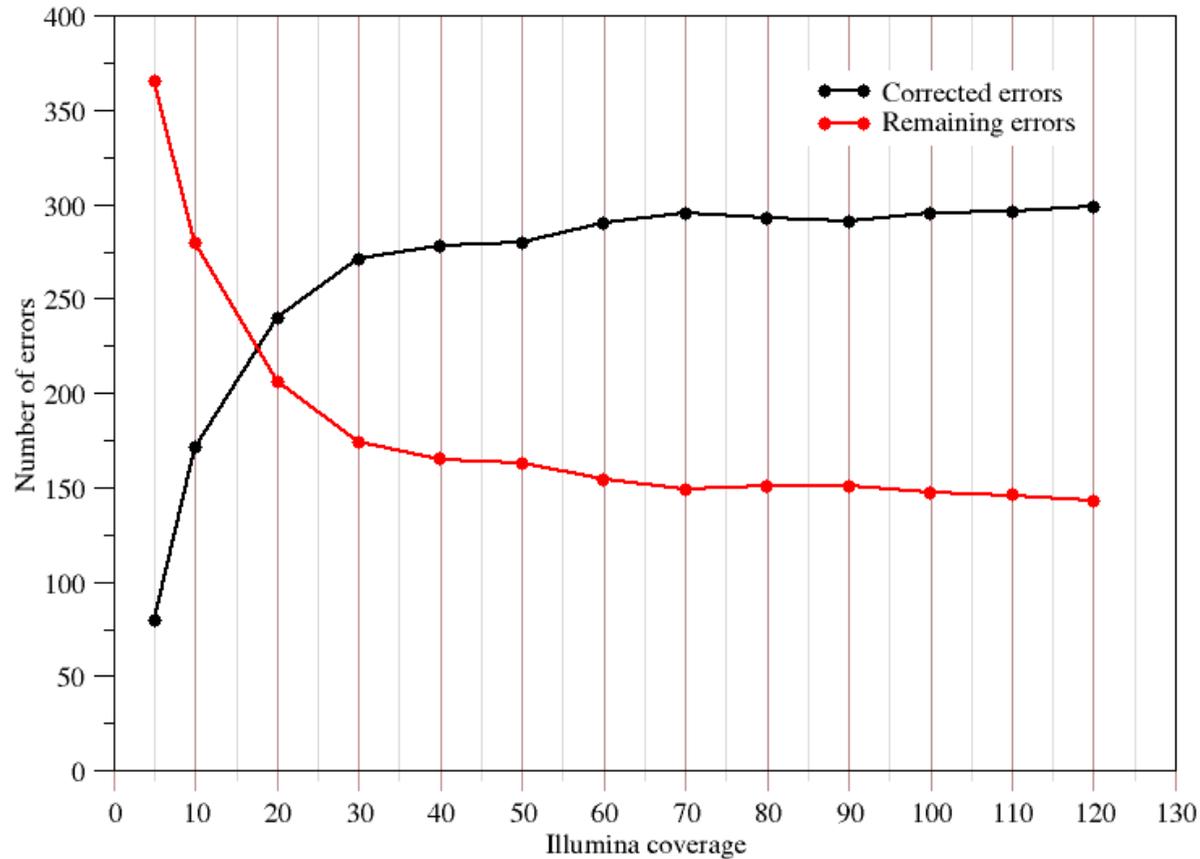
- Alignement des lectures illumina sur l'assemblage 454 en utilisant Soap (alignements gappés) : 2 mismatches et 3 gaps maximum
- Elimination des lectures alignées de façon non-unique
- Chaque différence est conservée si elle satisfait les critères suivants :
 - Elle n'est pas située dans les 5 premières et dernières bases de la lecture
 - La qualité de la base en question et des bases encadrantes est ≥ 20
 - Les séquences flanquantes ne sont pas des homopolymères
- Une différence est considérée comme une erreur de séquence si :
 - Elle est vue au moins par 3 lectures différentes
 - 70% des lectures alignées à cette position sont en accord
- Ces critères qualités entraînent une chute de la couverture

Séquençage de génomes

	Sanger	Unpaired + PE 454	unpaired + paired 454 with Illumina / Solexa GA1
Coverage	7.4X	25X	25X and 50X
Assembler	Arachne (Broad Institute)†	Newbler (454/Roche)†	Newbler (454 / Roche)†
# of contigs	173	119	119
Contigs N50 (Kb)†	39.0	58.2	58.2
# of scaffolds	2	10	10
Scaffolds N50 (Kb)†	2,200	1,000	1,000
Assembly size (% of reference)†	3.417Mb (95%)†	3.544 Mb (98%)†	3.544 Mb (98%)†
Mis-assemblies	0	0	0
# of errors	3,442	431 (1 erreur / 8Kb)†	163 (1 erreur / 22Kb)†
Substitutions	2,494	75	71
Insertions / Deletions	948	356	92

Séquençage de génomes

Correction des erreurs du 454



- A 50X, reste 163 erreurs :
 - 51 sont dues à des erreurs dans la séquence de référence ou à la présence de variations (cultures différentes).
 - 112 sont localisées dans les régions répétées (pas de couverture solexa) ou en extrémité de contigs.

Séquençage de génomes

Genomic DNA

↓
Roche/454 sequenced paired-end library
to a ~7x fragment size coverage (for 3Kb fragments)

↓
Add 454 unpaired data to a final 25x coverage

Newbler assembly

↓
Correct errors with ~50x Solexa/illumina
short reads data

High quality draft (< 10^{-4} error rate)

Plateforme détection de mutations par capture



- Laboratoire de Ressources Génomique : Gabòr Gyapay
- Laboratoire de Séquençage : Patrick Wincker
- Laboratoire d'Analyse Bioinformatique des Séquences : François Artiguenave, Vincent Meyer, Marc Wessner, Benjamin Noel

Plateforme détection de mutations

- Objectifs : détection de mutations
 - sur des ‘grands’ génomes (typiquement l’humain)
 - sur plusieurs individus en parallèle
 - pour un cout raisonnable
- Principe :
 - définir des régions d’intérêts sur ces ‘grands’ génomes de plusieurs mégabases
 - amplifier spécifiquement ces régions par capture
 - séquençage haut-débit
- Utilisation de puces Nimblegen pour la capture et séquençage en 454
- Quels types de projets ?
 - Maladies génétiques rares (dermatologie, neurologie, etc...).
 - Cancerologie.
 - Autres thématiques venant d’appels de proposition du Génoscope (analyse du génome humain et d’autres mammifères, etc...).



- **Alignement des lectures 454 sur le génome**

- ssaha2 (blat en cours d'intégration)

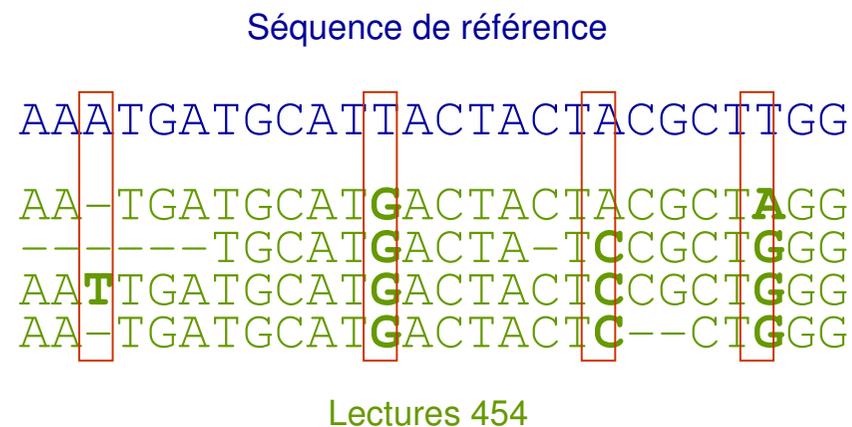
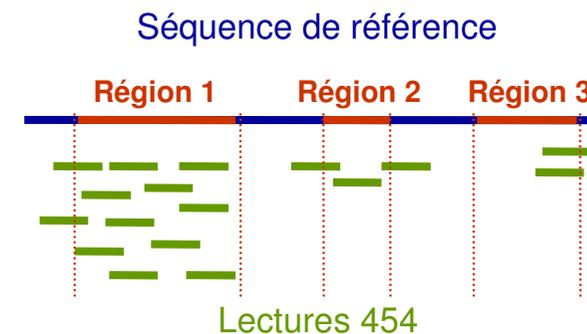
- **Evaluation de la qualité de couverture**

- % de couverture des régions d'intérêt
- % de lectures mappées incluses ou chevauchant les régions d'intérêt

- **Détection des variations**

- Utilisation de "gsMapper" fourni par Roche 454

- **Mise à disposition des résultats**



- **Détail d'un fichier résultat (HCDiff)**

```

Other Reads:

-----

>chr9 129980643 129980643 C T 3 1 4 4 100

Reads with Difference:
chr9 129980611+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGCTT-CTCOGTCOGCAGCCAGAACCTTGAGGCC 129980674
*****
FHJ73AY04JK9ZZ 74- TATTTCCAGAGTCACC-AGGCCCAAGAGCTGC-TTTCCTCOGTCOGCAGCCAGAAC 21
FHJ73AY04I4Q31 113+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCOGTCOGCAGCCAGAACCTTGAGGCC 176
FJXEIA402G769E 92+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCOGTCOGCAGCCAGAACCTTGAGGCC 155
FJXEIA402FOMZA 27+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCOGTCOGCAGCCAGAACCTTGAGGCC 90
*****

Other Reads:
  
```

- positionnement sur le génome de référence

● Détail d'un fichier "HCDiff"

```

Other Reads:

-----

>chr9    129980643      129980643      C      T      3      1      4      4      100

Reads with Difference:
chr9      129980611+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGCTT-CTCGTTCGGCAGCCAGAACCTTGAGGCC 129980674
                        *****

FHJ73AY04JK9ZZ      74- TATTTCCAGAGTCACC-AGGCCCAAGAGCTGC-TTTCCTCGTTCGGCAGCCAGAAC      21
FHJ73AY04I4Q31     113+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCGTTCGGCAGCCAGAACCTTGAGGCC 176
FJXEIA402G769E     92+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCGTTCGGCAGCCAGAACCTTGAGGCC 155
FJXEIA402FOMZA     27+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCGTTCGGCAGCCAGAACCTTGAGGCC 90
                        *****

Other Reads:

```

- positionnement sur le génome de référence
- **séquence de référence**

● Détail d'un fichier "HCDiff"

```

Other Reads :
-----
>chr9  129980643      129980643      C      T      3      1      4      4      100

Reads with Difference:
chr9          129980611+  TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGCTT-CTCGTCCGCAGCCAGAACCTTGAGGCC 129980674
                                     *****
FHJ73AY04JK9ZZ      74-  TATTTCCAGAGTCACC-AGGCCCAAGAGCTGC-TTCTCCGTCGGCAGCCAGAAC      21
FHJ73AY04I4Q31     113+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTCTCCGTCGGCAGCCAGAACCTTGAGGCC 176
FJXEIA402G769E     92+  TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTCTCCGTCGGCAGCCAGAACCTTGAGGCC 155
FJXEIA402FOMZA     27+  TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTCTCCGTCGGCAGCCAGAACCTTGAGGCC 90
                                     *****

Other Reads :
  
```

- positionnement sur le génome de référence
- séquence de référence
- **séquences des lectures présentant des différences**

● Détail d'un fichier "HCDiff"

```

Other Reads:
-----
>chr9  129980643      129980643      C      T      3      1      4      4      100

Reads with Difference:
chr9          129980611+  TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGCGTT-CTCGTTCGTCAGCCAGAACCTTGAGGCC 129980674
                                     *****
FHJ73AY04JK9ZZ      74- TATTTCCAGAGTCACC-AGGCCCAAGAGCTGC-TTTCCTCGTTCGTCAGCCAGAAC      21
FHJ73AY04I4Q31     113+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCGTTCGTCAGCCAGAACCTTGAGGCC 176
FJXEIA402G769E     92+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCGTTCGTCAGCCAGAACCTTGAGGCC 155
FJXEIA402FOMZA     27+ TATTTCCAGAGTC-CCCAGGCCCAAGAGCTGC-TTTCCTCGTTCGTCAGCCAGAACCTTGAGGCC 90
                                     *****

Other Reads:

```

- positionnement sur le génome de référence
- séquence de référence
- séquences des lectures présentant une variation

Rechercher une position ou un nom

Pour réaliser une recherche, vous pouvez utiliser une position (chr:début-fin), un nom de gène ou un identifiant de variation. Le caractère "*" est autorisé

Exemples :

[chr18:27701000..27701650](#) région du chromosome 18 comprise entre les positions 27701000 et 27701650

[ENSG00000189461](#) identifiant d'un gène *Ensembl*

[C5076](#) identifiant d'une variation du projet YT

A propos du projet YT

L'objectif de ce projet est d'identifier de nouveaux gènes suppresseurs de tumeurs impliqués dans les cancers de la vessie par l'identification de mutations. Plusieurs stratégies ont été employées pour sélectionner de potentiels gènes suppresseurs de tumeurs:

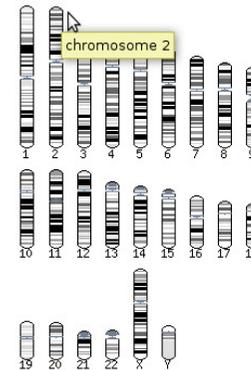
- 1 - L'analyse des échantillons tumoraux de vessie par Hybridation Génomique Comparative (puces CGH) a permis de mettre en évidence des régions génomiques fréquemment perdues dans ces cancers. Les gènes et les microARNs localisés dans ces régions de perte ont été sélectionnés pour séquençage.
- 2 - Liste des gènes connus par la base de données "Cancer Gene Census" comme ayant des mutations impliqués dans la progression tumorale. Seuls les gènes avec des mutations somatiques et non mutés dans les lignées germinatives ont été retenus.
- 3 - Liste des gènes connus pour être mutés dans les lignées cellulaires.
- 4 - Liste des gènes connus pour être localisés dans des régions génomiques sous régulation épigénétique dans les cancers de la vessie.
- 5 - Liste des gènes constitutifs des voies de signalisation impliquées dans la progression tumorale.

Une liste de 1251 gènes (y compris microARN) a donc été constituée pour ce projet de recherche de mutations dans les cancers de la vessie. Cela équivaut à un total de 3,943,922 nucléotides (5,302,487 nucléotides pour le comptage Nimblegen)

Ce projet de séquençage a été réalisé pour 8 échantillons: 4 échantillons tumoraux & 4 échantillons normaux appariés (provenant des mêmes patients). Le séquençage en parallèle sur des échantillons tumoraux et normaux appariés a été effectué afin de faciliter l'analyse des polymorphismes identifiés par le séquençage et notamment de pouvoir différencier les SNPs des mutations. Les échantillons sélectionnés remplissaient les conditions suivantes: (i) ADN génomique extrait en quantité suffisante et de qualité et (ii) échantillons contenant les régions CGH récurrentes de pertes.

Institut Curie : [Laboratoire d'Oncologie moléculaire](#)

Explorer un chromosome



Statistiques YT

2244442 séquences chevauchent les régions ciblées (soit 61% de l'ensemble des lectures)

% de régions couvertes par des lectures : 90%

% de régions entièrement couvertes : 79%

Couverture moyenne des régions : 12.42 X

(Valeurs moyennes sur 6 individus)

[Bilan détaillé individu par individu...](#)

Genome Browser incluant la représentation des SNPs


Centre National de Séquençage
PLATEFORME DE DETECTION DE MUTATIONS

[Genoscope](#) | [Plateforme](#) | [Projet_YT](#) | [Navigateur](#) | [Analyse](#) | [Téléchargements](#) | [Contact](#)

Vue de 1.001 kbp depuis chr11, positions 297,814 à 298,814

Instructions
 Vous pouvez faire une recherche en utilisant un nom de séquence, un nom de gène, un locus, ou un autre référentiel. Le caractère spécial * est autorisé. Pour vous recentrer sur un emplacement, cliquez sur la règle. Utilisez les boutons Défil./Zoom pour changer l'échelle et la position. Pour sauvegarder cette vue, ajoutez ce lien à vos favoris.

Exemples: chr1:554450-555000, chr18:27701000..27701650, chrY:11928220..11928330.

[Cacher l'en-tête] [Ajouter cet affichage à vos favoris] [Lien vers une image de cet affichage] [Image haute qualité pour les publications] [Aide] [Remise à zéro]

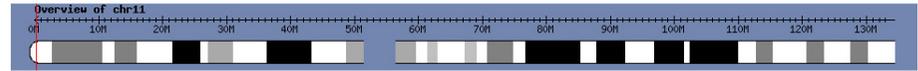
Chercher
 Référentiel ou Région:

Sorties, recherches et autres opérations:
 Sortie Decorated FASTA File

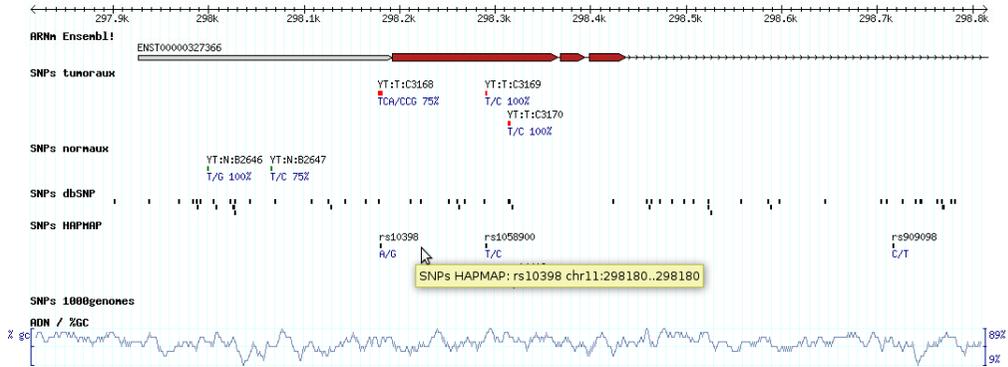
Source de données
 Projet YT

Défil./Zoom: Inversion

Aperçu



Détails



Supprimer le surbrassage

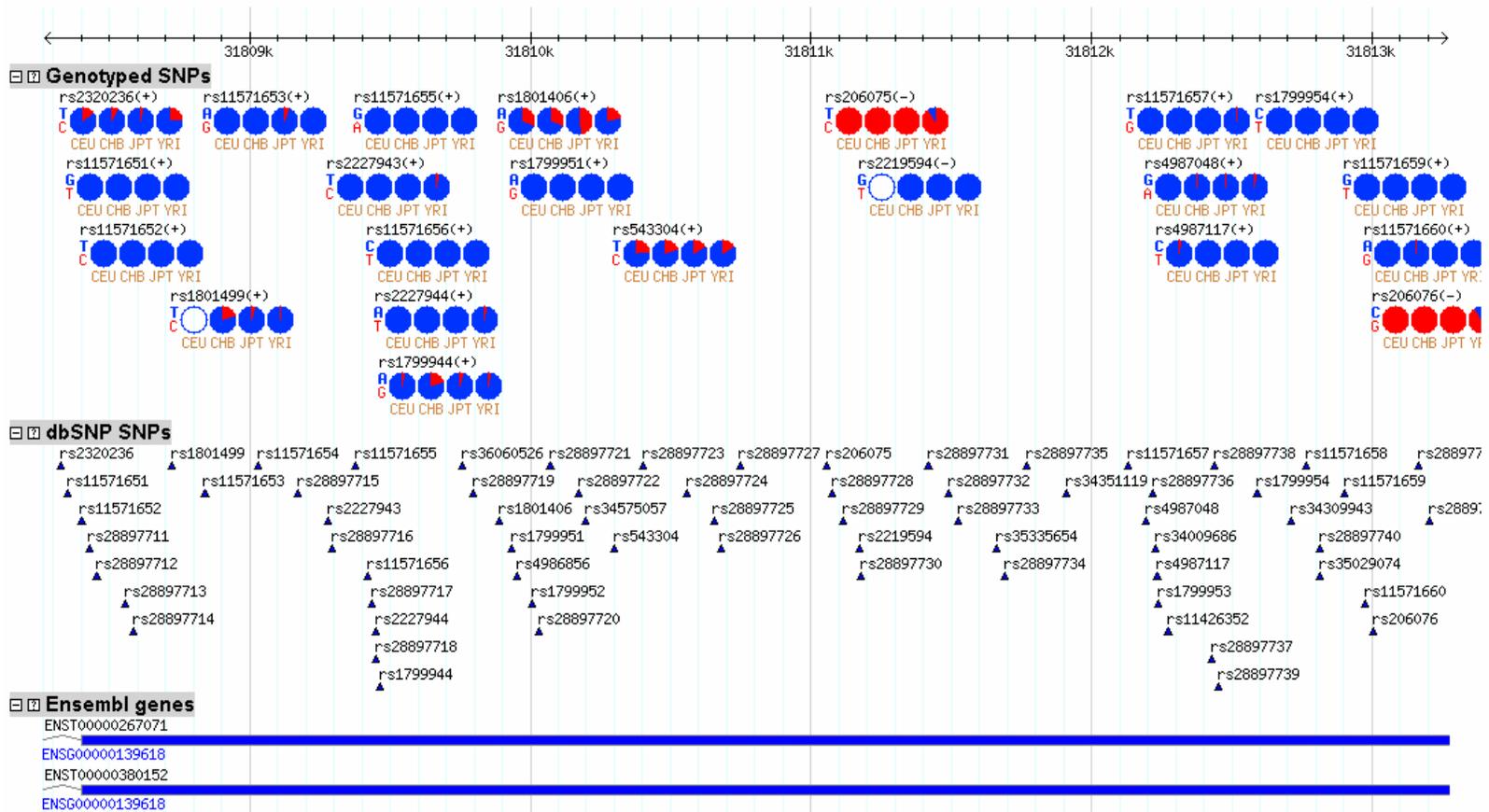
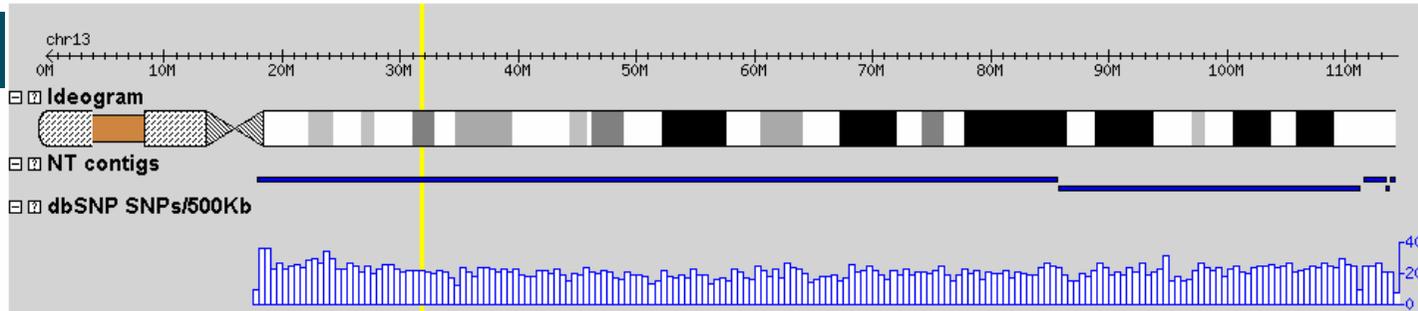
Pistes
 Préférences d'affichage
 Ajouter vos propres pistes

© Genoscope - Centre National de Séquençage - 2 rue Gaston Crémieux CP5706 91057 Evry cedex - Tél: (+33) 0 1 60 87 25 00 - Fax: (+33) 0 1 60 87 25 14 - Exécution: 14/04/2009 - 13:21:37

NB: Cette page utilise un cookie pour sauver et restituer les informations de configuration. Vos informations ne sont pas partagées.
 Generic genome browser version 1.68



Liens dynamique vers le site Hapmap



Interface de requête



Centre National de Séquençage

PLATEFORME DE DETECTION DE MUTATIONS

Genoscope | Plateforme | Projet_YT | Navigateur | Analyse | Téléchargements | Contact

OUTIL D'ANALYSE DES VARIATIONS

PROJET YT

INDIVIDUS

Tous | Aucun

Union Intersection B C D E F H annotation

CHROMOSOMES

Tous | Aucun

Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8
 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16
 Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 ChrX ChrY

VARIATIONS

% de lectures présentant la variation > à HCDiff ALLDiff Trier en fonction du pourcentage

Tous | Aucun

Absentes des bases de données publiques : dbSNP I1apMap 1000genomes

Tous | Aucun

Absente chez les individus : B C D E F H annotation

LOCALISATION

Tous | Aucun

Intergénique mRNA 5'UTR' Exon Intron 3'UTR'

ANNOTATION

Hétérozygote Homozygote

Non Synonyme Synonyme

Format de sortie : .txt .bed .gff

OUTIL D'ANALYSE DES VARIATIONS

PROJET YT

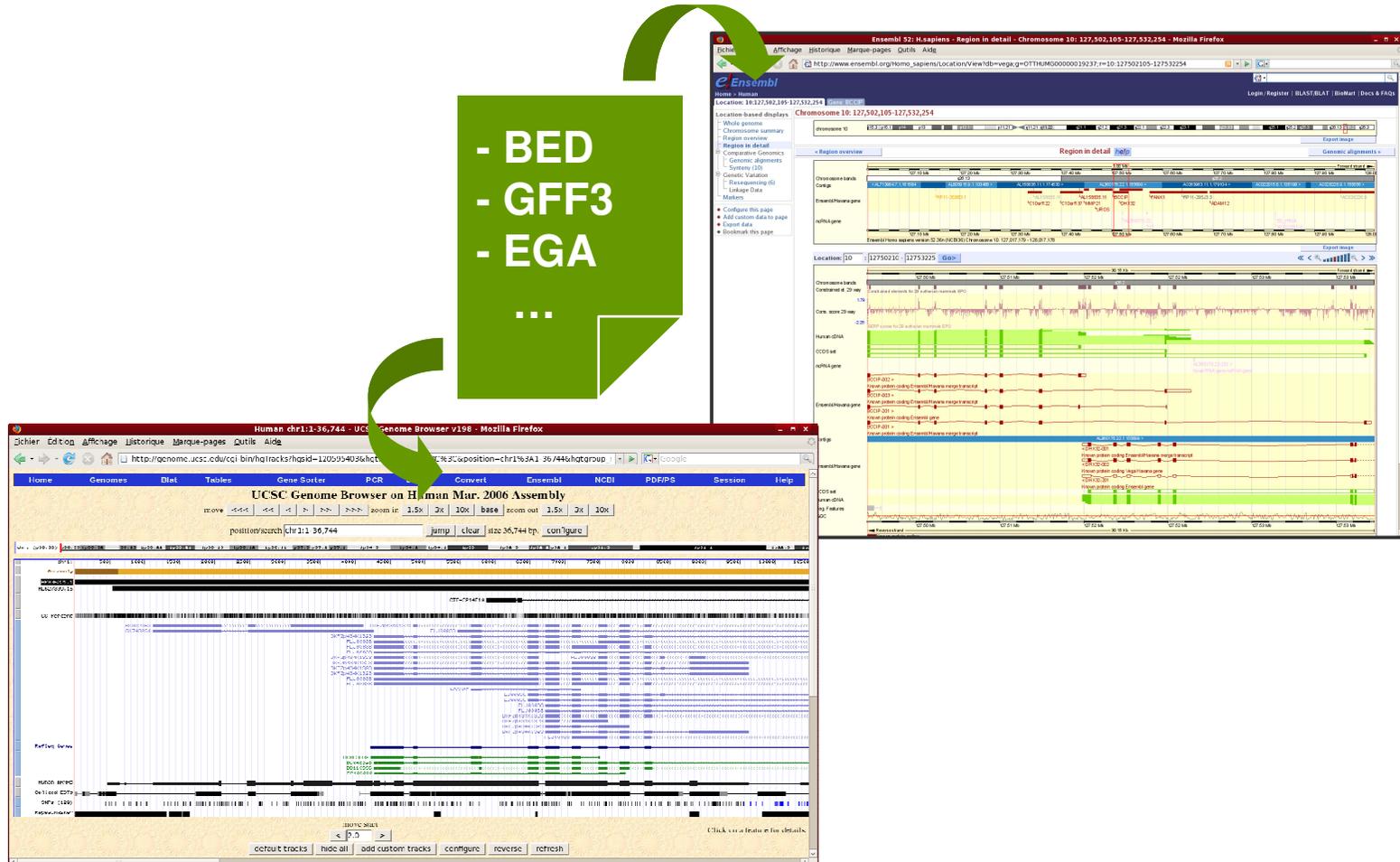
23 variations observées répondent à vos critères.

[Pour télécharger le fichier résultat : Enregistrer la cible du lien sous...](#)

Individu	Ref.	Chr	Debut	Fin	Variation	Nb total Lectures	% lectures avec var.	Type de Localisation ▼	Identifiant de Localisation	Annotation	dbSNP	HapMap	1000 genomes
B	B11	chr1	4275	4275	T/C	2+5=12	58 %	CDS	ENST00000326632	SHIFT			
C	C31	chr1	4275	4275	T/C	2+5=12	58 %	CDS	ENST00000326632	SHIFT			
C	C51	chr1	1675900	1675900	G/T	1+4=5	100 %	intergénique					
C	C32	chr1	2435266	2435266	A/G	3+1=4	100 %	intergénique					
B	B2	chr1	2435266	2435266	A/G	3+1=4	100 %	intergénique					
B	B1	chr1	1675900	1675900	G/T	1+4=5	100 %	intergénique					
B	B20	chr1	2960694	2960694	G/T	1+4=5	100 %	intergénique					T/A
B	B19	chr1	3693803	3693803	G/T	1+4=5	100 %	intergénique			-T		
B	B4	chr1	2442304	2442304	A/G	4+2=8	75 %	intergénique					
B	B5	chr1	2442429	2442429	T/C	2+3=7	71 %	intergénique					
C	C33	chr1	2442217	2442217	G/A	3+2=7	71 %	intergénique					
C	C25	chr1	2442429	2442429	T/C	2+3=7	71 %	intergénique					
B	B3	chr1	2442217	2442217	G/A	3+2=7	71 %	intergénique					
C	C36	chr1	15767642	15767642	G/A	5+3=13	62 %	intergénique					
B	B6	chr1	15767642	15767642	G/A	5+3=13	62 %	intergénique					
C	C10	chr1	32571790	32571790	T/C	2+5=12	58 %	intergénique					
B	B17	chr10	62765	62765	T/C	2+5=12	58 %	intergénique				C/T	
B	B16	chr1	38125497	38125497	T/C	2+5=12	58 %	intergénique					
C	C37	chr10	62765	62765	T/C	2+5=12	58 %	intergénique				C/T	
B	B15	chr1	38125497	38125497	T/C	2+5=12	58 %	intergénique					
C	C27	chr1	15768020	15768020	G/A	5+3=20	40 %	intergénique					
B	B7	chr1	15768020	15768020	G/A	5+3=20	40 %	intergénique					
B	B8	chr1	15769345	15769345	A/G	1+3=10	40 %	intergénique					

Mise à disposition de fichiers permettant de visualiser les données du projet dans un browser génomique ...

- BED
- GFF3
- EGA
...



	Disponible	En projet
Mapping		
- Assuré par ssaha2 (1journée/run)	○	
- Implémentation de BLAT pour un gain de temps de calcul (1h/run)		○
Résultats		
- Information sur la qualité de la couverture des zones étudiées (mutastat)	○	
- Détection des mutations (logiciel constructeur : GSMapper)	○	
- Évaluation de la qualité des résultats (calibration via dbSNP)		○
- Approches “paired-end” et “structure variation”		○
- Développement de nouvelles méthodes de scoring		○
Interprétation		
- Comparaison de profils de mutations	○	
- Fonctionnalités de croisement des données (Filtres dbSNP/1000 génomes/HapMap...)	○	
- Visualisation des résultats (GenomeBrowser)	○	
- Caractérisation fonctionnelle de la mutation: (Caractérisation de la zone, modification du cadre de lecture, modification acide aminé)	○	





Méthodes de détection de mutations

- ✓ Algorithmes qui tiennent compte des spécificité des données issues des NTS.
- ✓ Algorithmes existants :
 - ✓ SSAHA_pileup (Sanger), version adaptée de SSAHA_snp
 - ✓ GigaBayes (Boston College), version adaptée de PolyBayes



Neighbourhood Quality Standard (NQS)

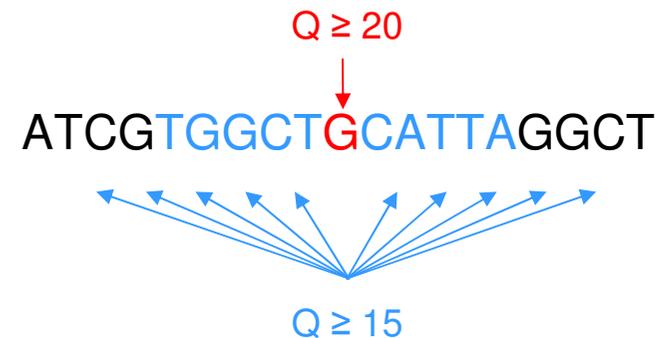
Altshuler, D. *et al.* An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407, 513-516 (2000)

Méthode basée sur les scores qualités attribués à chaque base des lectures.

Le NQS a pour objectif de diminuer l'impact du taux d'erreurs du 'base calling' sur l'identification des sites polymorphes

Une base est marquée NQS si :

- ✓ son score qualité $Q \geq 20$
- ✓ les 5 bases de chaque côté ont chacune un score $Q \geq 15$





Méthodes de détection de mutations

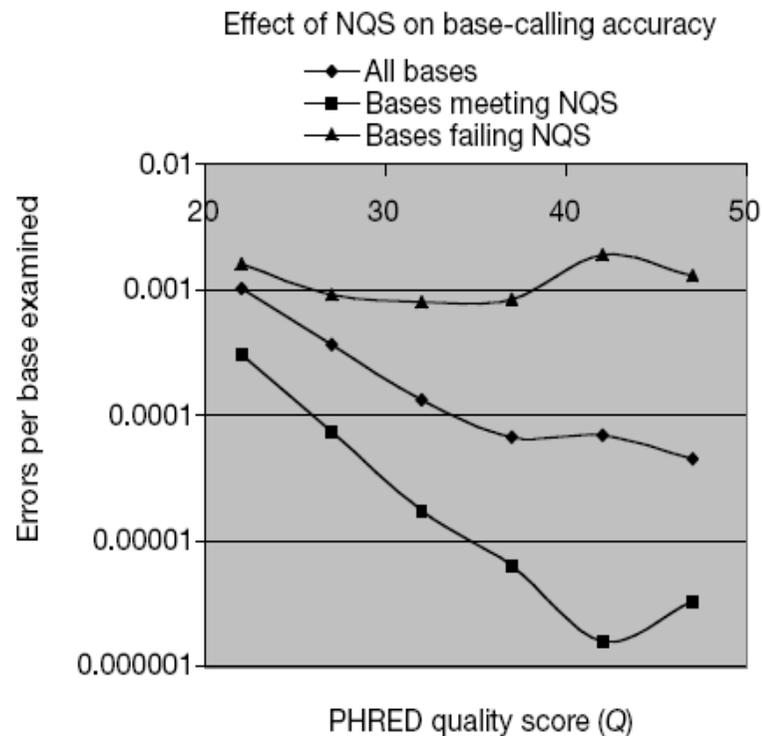
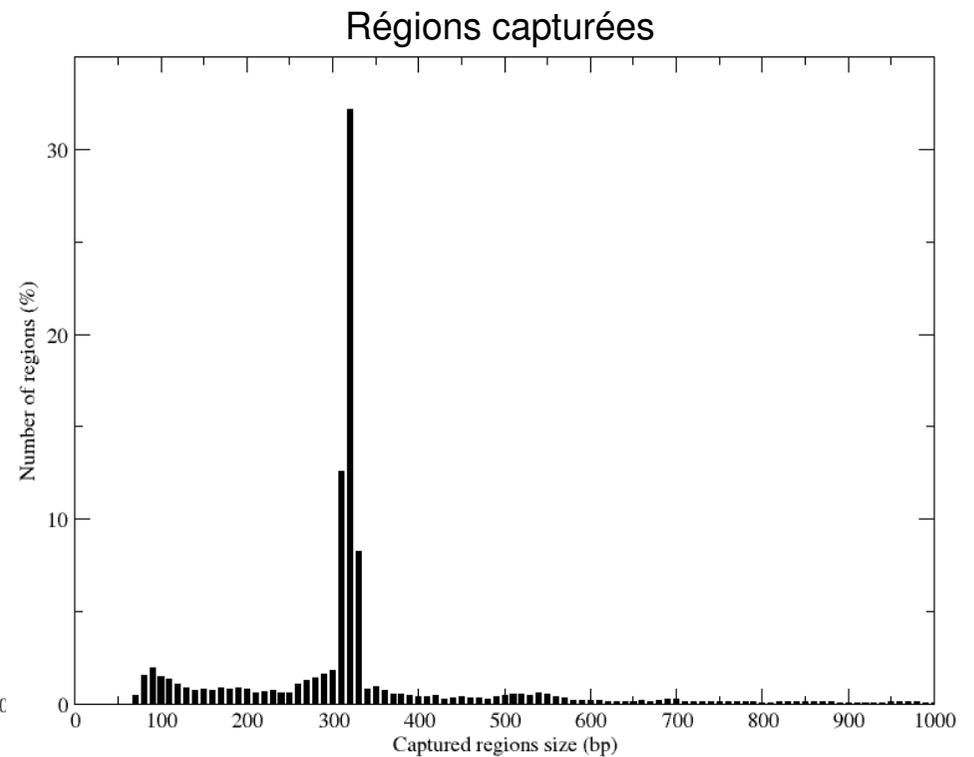
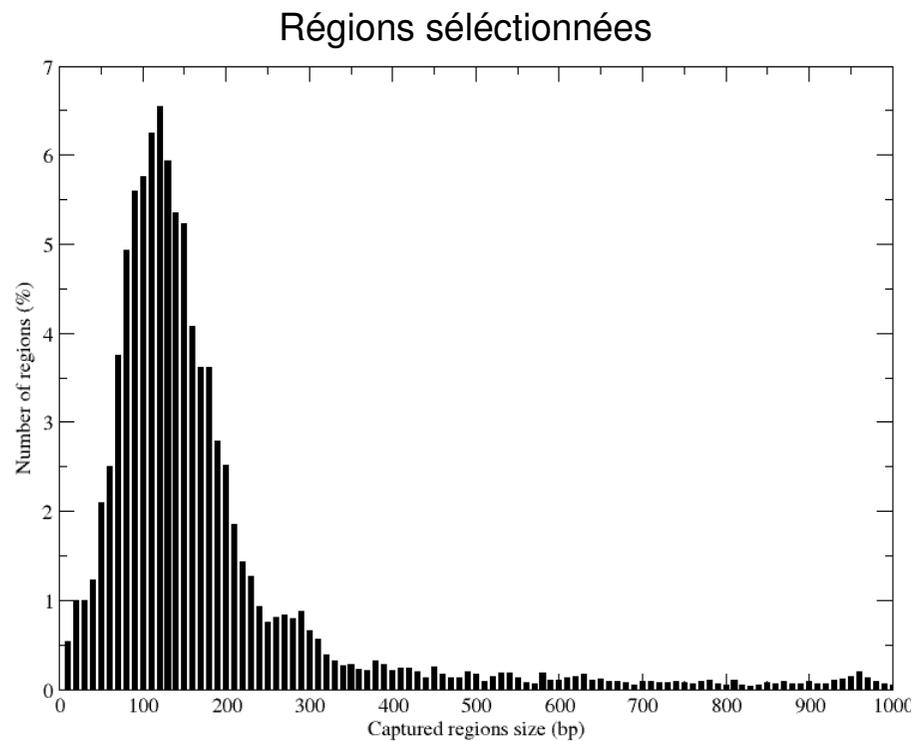


Figure 1 Impact of quality criteria on error rates. Data are plotted according to the PHRED Q score of each base^{8,9} and reported in bins of five PHRED Q units; only base substitution errors were counted. As previously reported⁹, overall PHRED scores accurately predict the observed rates of base-calling errors; however, bases meeting the NQS display substantially lower error rates than are predicted by their PHRED scores. Although the effect is proportionally greatest for high PHRED scores, the bulk of errors avoided are found in bases with lower PHRED scores (that is, those with the highest error rates).

Altshuler, D. *et al.* An SNP map of the human genome generated by reduced representation shotgun sequencing. *Nature* 407, 513-516 (2000)

Plateforme détection de mutations

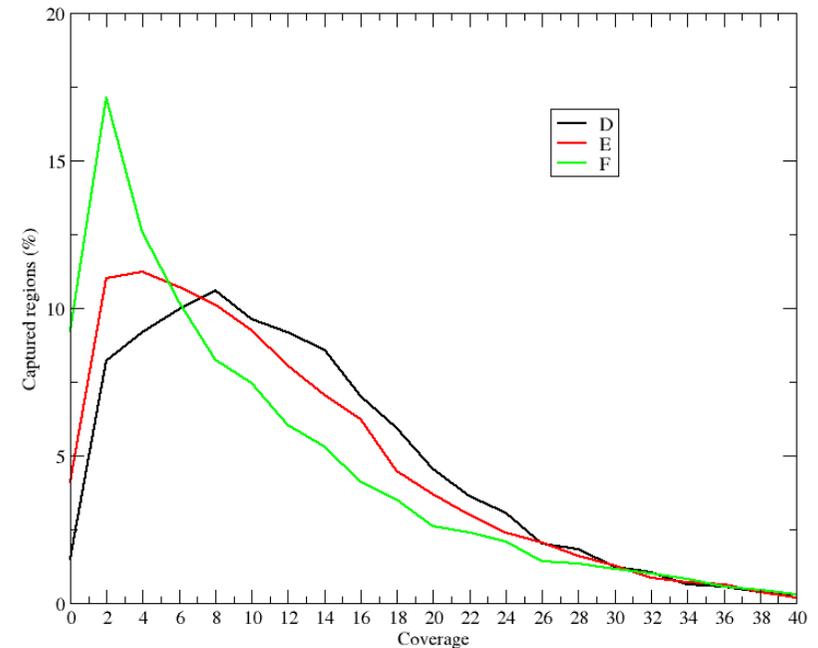
- ✓ Projet pilote : sélection de 1.251 gènes , 13.315 exons , taille cumulée d'environ 4 Mb
- ✓ 8 échantillons avec 1 run GSFLX par échantillon (soit ~ 100Mb)
- ✓ 13.315 régions ciblées : 3,97Mb (moyenne de 300pb)
- ✓ Après passage chez NimbleGen : 13.944 régions ; 5,6Mb (moyenne de 400pb)



Plateforme détection de mutations

	B	C	D	E	F	H
Couverture initiale	42,9	53,3	35,1	35,1	41,1	29,6
Couverture moyenne	13,9	15,8	12,7	11,5	10,5	10,1
Couverture minimale	0	0	0	0	0	0
Couverture maximale	80,7	102,2	102,1	113,0	111,0	86,0
# régions couvertes à 10X	7.026 (53%) [†]	7.985 (60%) [†]	7.123 (54%) [†]	5.886 (44%) [†]	4.097 (31%) [†]	5.502 (41%) [†]

Avec >30X initialement, on ne couvre qu'environ 50% des régions avec une couverture supérieure à 10X



Plateforme détection de mutations

- ✓ Alignement des lectures provenant des 8 échantillons sur le génome humain

