



**cgfb**

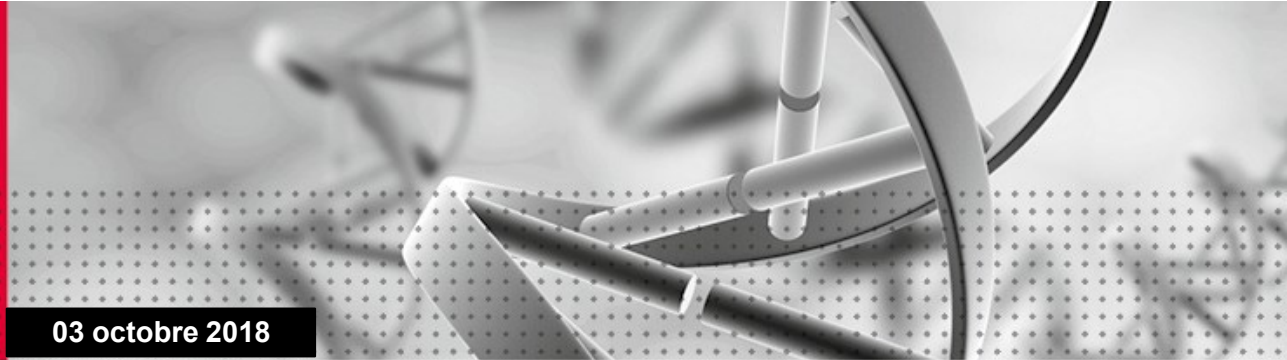
GÉNOME TRANSCRIPTOME

03 octobre 2018

# TruSeq NE – Application

*I. Lesur Kupin & F. Salin*

Colloque EPGV  
3-5  
Octobre 2018





➔ R&D et prestations en séquençage - **génotypage**

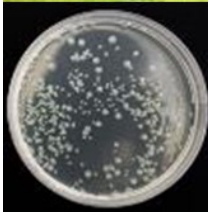
**SÉQUENCER**

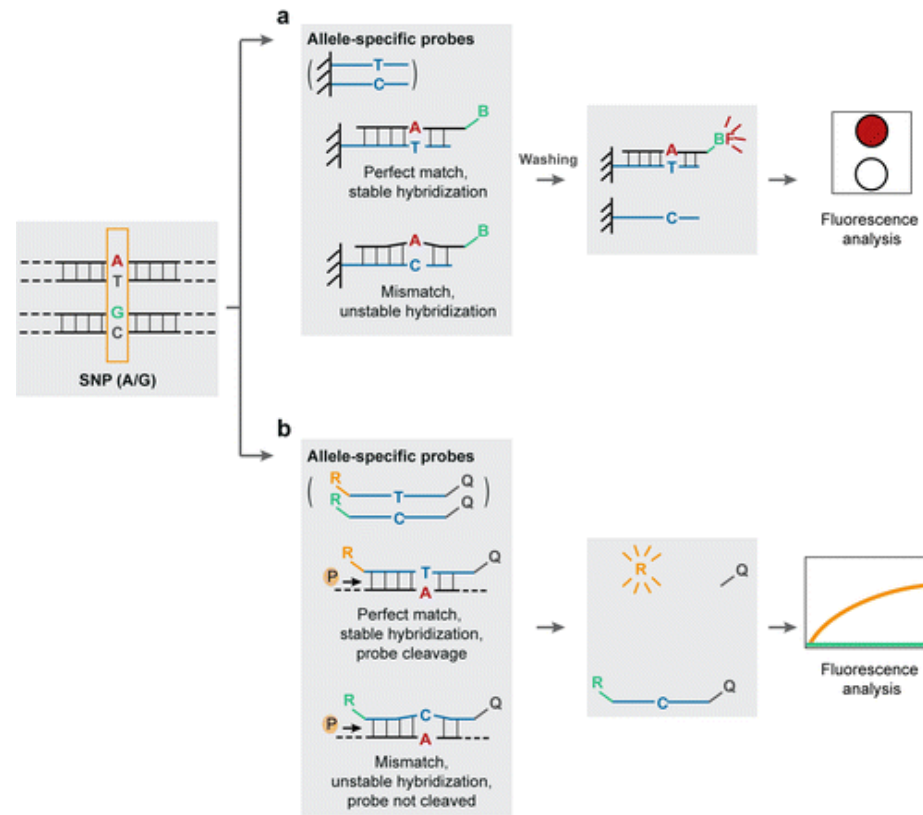
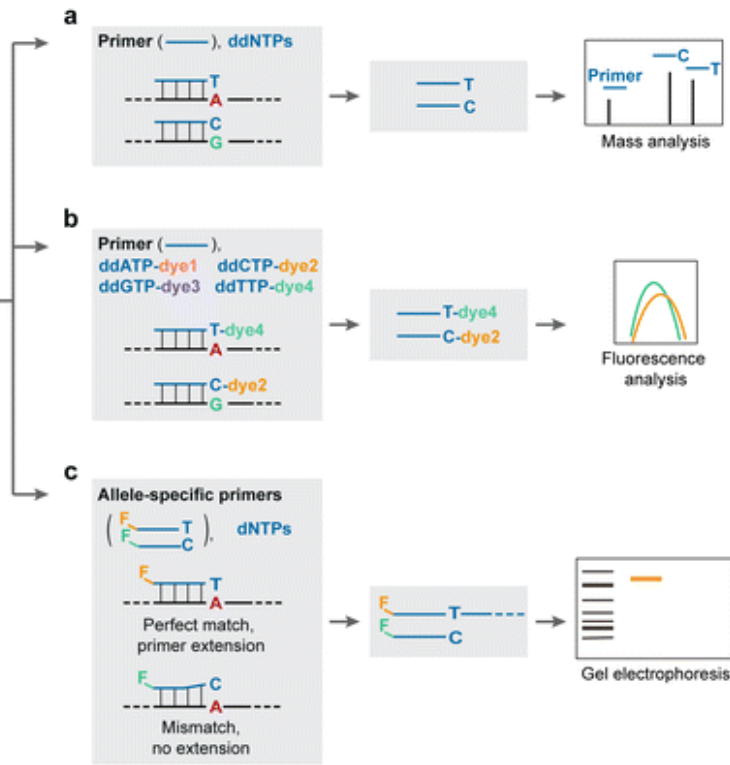
**GÉNOTYPER**

**ANALYSER**

- Conseil-suivi de projets
- Production de données
- Analyse de données

➔ Formation et mise à disposition





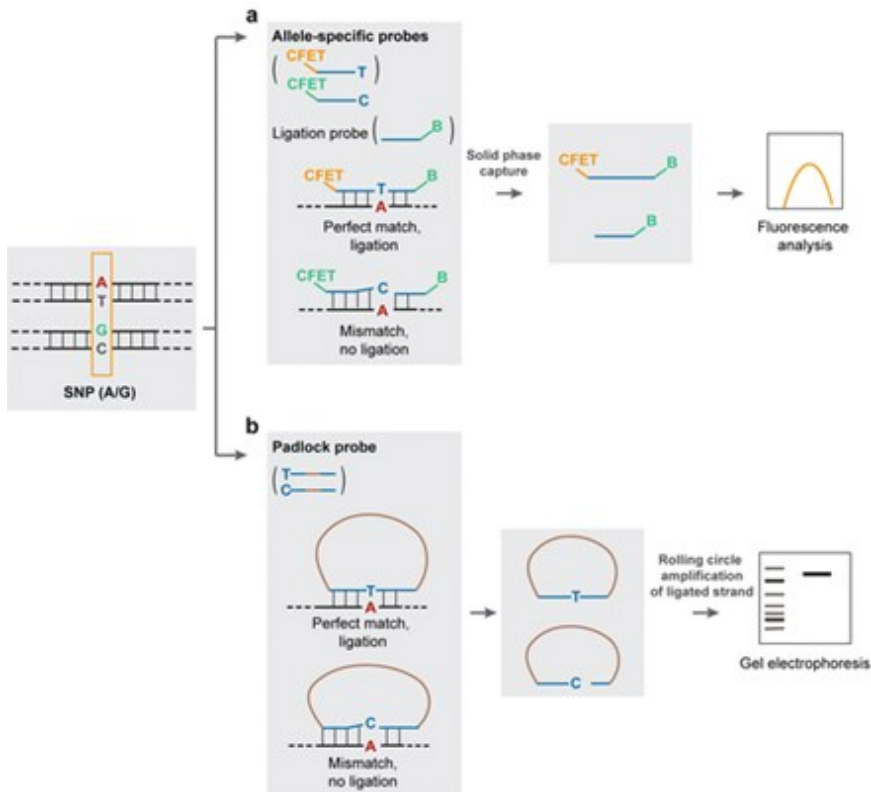
Kim S, Misra A. 2007. Annu. Rev. Biomed. Eng. 9:289–320

Kim S, Misra A. 2007. Annu. Rev. Biomed. Eng. 9:289–320

Primer extension

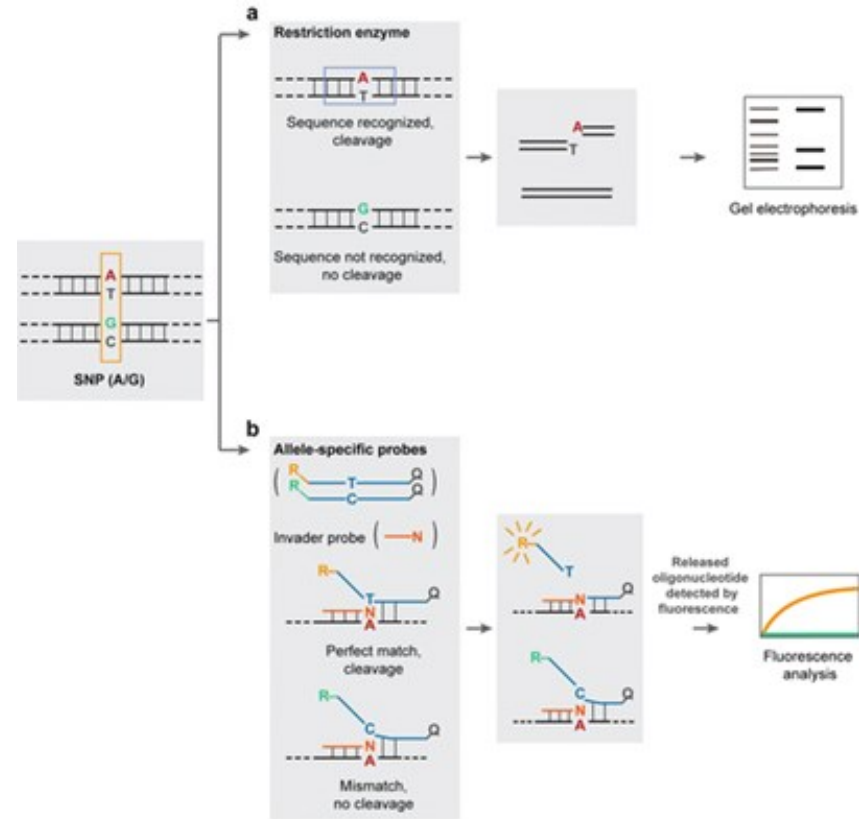
Hybridization

### Ligation



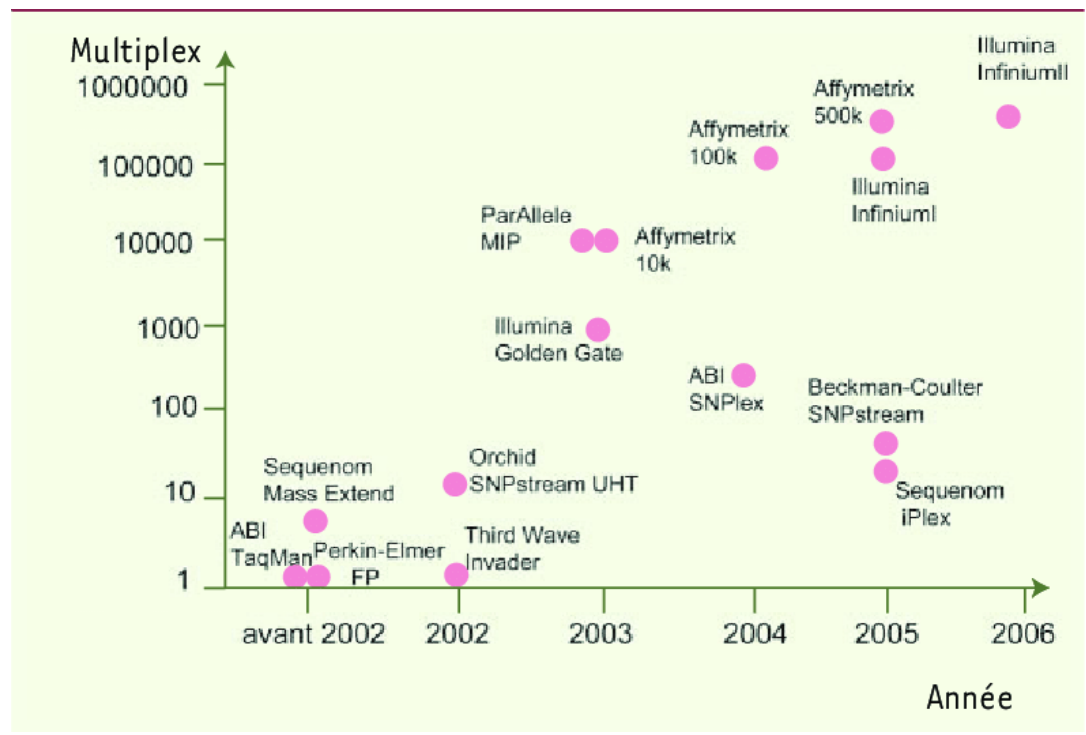
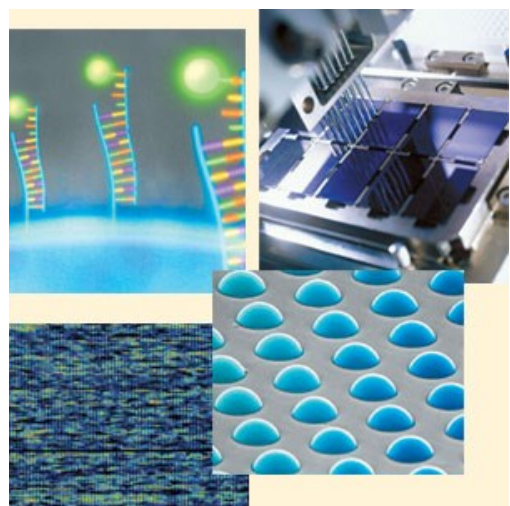
Kim S, Misra A. 2007. Annu. Rev. Biomed. Eng. 9:289-320

### Clivage enzymatique



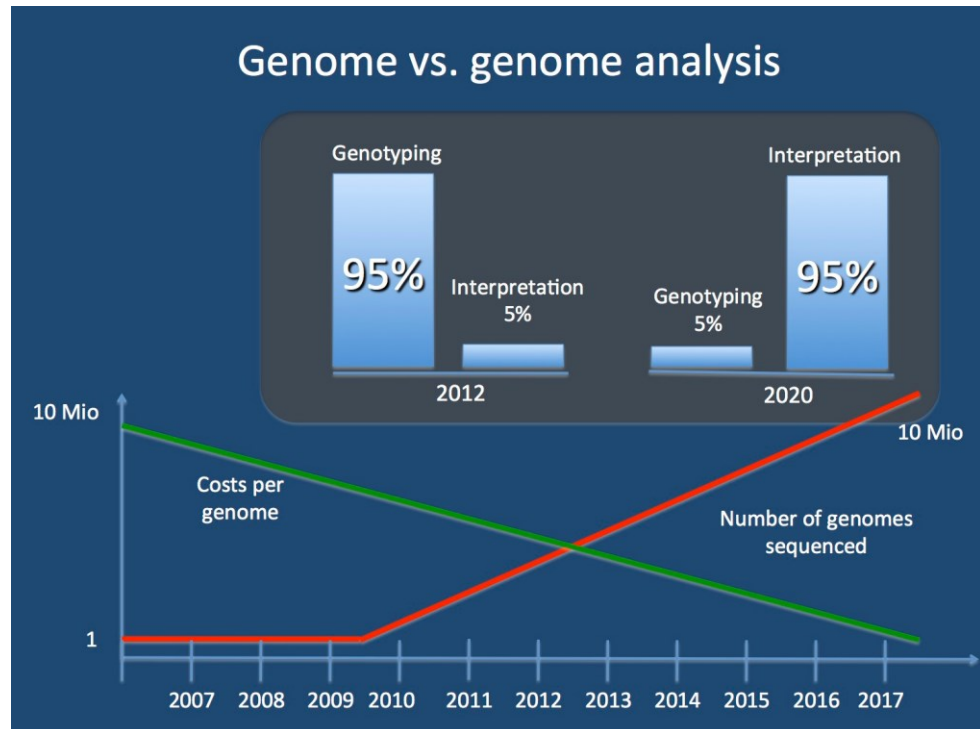
Kim S, Misra A. 2007. Annu. Rev. Biomed. Eng. 9:289-320

# Plateforme Génome-Transcriptome de Bordeaux



Critères de choix de la technologie :

- Nombre de SNP
- Nombre d'individus
- Espèce modèle ou pas (existence de puces commerciales)
- Le **prix**



## True seq NE

The TruSeq Genotype N<sub>e</sub> Kit is a low-cost, targeted solution for genotyping by sequencing. The kit offers several advantages to support various applications, including parentage studies, purity testing, and marker-assisted selection.

- Detection of up to 5000 targets in a single panel
- Multiplexing capabilities to process up to 384 samples per run
- Fully supported, automation-friendly workflow that generates data in less than 2 days
- Custom design and panel optimization service with Illumina Concierge services

System Compatibility	<a href="#">MiniSeq</a> , <a href="#">MiSeq</a> , <a href="#">NextSeq 500</a> , <a href="#">NextSeq 550</a>
Species Category	Plant, Mammalian, Shrimp, Rice, Chicken, Soybean, Equine, Maize, Ovine, Porcine, Canine, Bovine, Rat, Mouse
Species Details	Any non-human animal or plant
Automation Capability	Liquid Handling Robots
Variant Class	Germline Variants, Insertions-Deletions (indels), Single Nucleotide Polymorphisms (SNPs)
Method	<a href="#">Amplicon Sequencing</a> , Custom Sequencing, <a href="#">Genotyping by Sequencing</a> , <a href="#">Targeted DNA Sequencing</a>
Assay Time	~6 hours total assay time
Hands-On Time	3 hours
Input Quantity	50 ng DNA
Content Specifications	Design custom probes to sequence genomic regions of interest. Content range: 16-5,000 targets
Mechanism of Action	Probe hybridization, extension-ligation, and PCR
Multiplexing	Up to 384 uniquely indexed samples may be pooled and sequenced together.
Nucleic Acid Type	DNA
7 Technology	Sequencing

## True seq NE

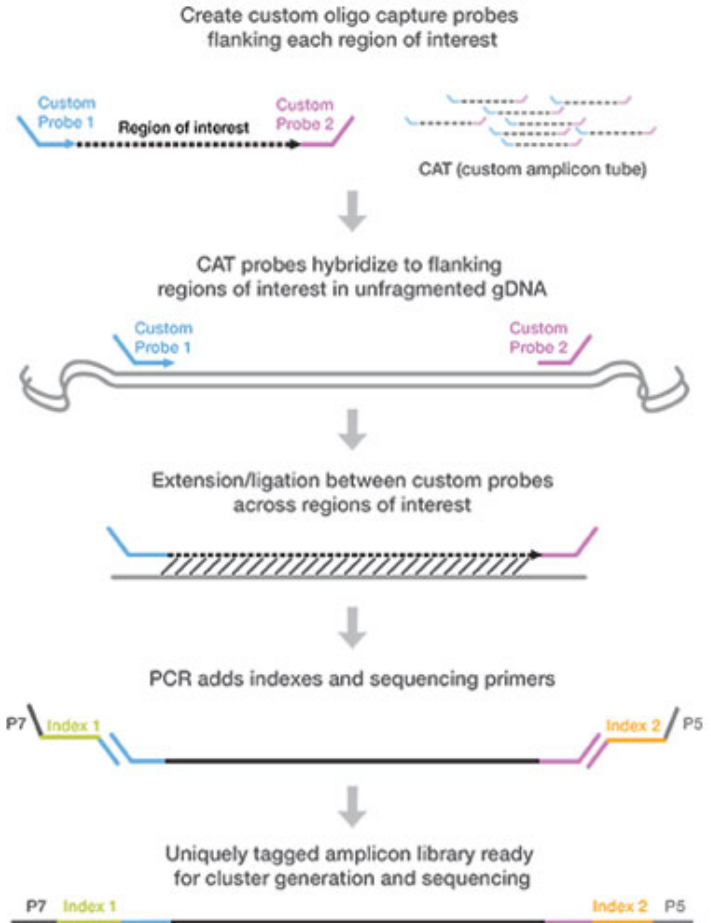
- 1 Quantify and Dilute DNA**  
 Hands-on: 10 minutes  
 Total: 10 minutes  
 Reagents: gDNA, RS1, SS1
- 2 Hybridize Oligo Pool**  
 Hands-on: 15 minutes  
 Total: 1 hour 50 minutes  
 Reagents: gDNA, CAT, OHS2, RS1  
 Plate: HYP
- 3 Remove Unbound Oligos**  
 Hands-on: 20 minutes  
 Total: 20 minutes  
 Reagents: SW1, SPB, Fresh 60% EtOH  
 Plate: HYP
- 4 Extend and Ligate Bound Oligos**  
 Hands-on: 20 minutes  
 Total: 1 hour 10 minutes  
 Reagents: ELB/ELE mixture  
 Plate: HYP
- 6 Amplify Libraries**  
 Hands-on: 30 minutes  
 Total: 2 hours—2 hours 15 minutes  
 Reagents: EDP/EMM mixture, i5 adapters, i7 adapters  
 Plate: HYP
- 7 Clean Up Libraries**  
 Hands-on: 20 minutes  
 Total: 30 minutes  
 Reagents: RSB, SPB, Fresh 80% EtOH  
 Plate: CLP, LNP
- 8 Normalize Libraries**  
 Hands-on: 30 minutes  
 Total: 50 minutes  
 Reagents: LNA1, LNB1 LNW1, LNS2, 0.1 N NaOH  
 Plate: SGP
- 9 Pool Libraries**  
 Hands-on: 5 minutes  
 Total: 5 minutes  
 Reagents: HT1  
 Tube: PAL

Safe Stopping Point

Safe Stopping Point

Safe Stopping Point

Safe Stopping Point



672 échantillons et 5000 targets  
 0.01€ le point



03 octobre 2018

# TruSeq NE – Application

*I. Lesur Kupin & F. Salin*



**cgfb**  
CENTRE BORDEAUX  
GÉNOMIQUE FONCTIONNELLE

# Le Projet

INRA Nancy  
UMR SILVA  
UEFL  
**Silva** UMR

INRA Paris  
URGV



# H<sub>2</sub>Oak

(O. Brendel)  
2014 - 2019

INRA Bordeaux  
UMR Biogeco



ONF



# Le Projet

**CONTEXTE:**

Changements climatiques

- Intensification des périodes de sécheresse
- conséquences sur les forêts françaises ?

**MODÈLE:**

le chêne

- Chêne pédonculé (*Q. robur*): tolérant à l'envoyage racinaire
- Chêne sessile (*Q. petraea*): tolérant à la sécheresse → utilise l'eau plus efficacement

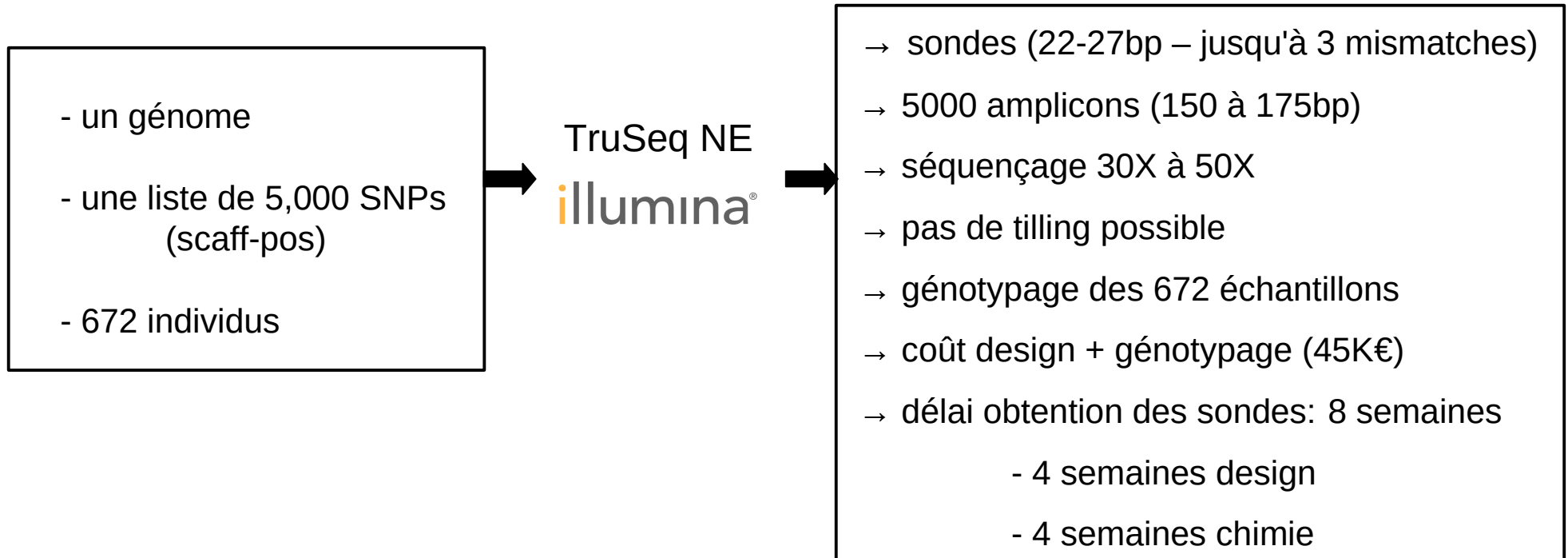
**QUESTION:**

Quantification du potentiel adaptatif des arbres pour l'efficacité d'utilisation de l'eau (WUE)

Polymorphisme des gènes candidats ↔ variation WUE ?

**TECHNOLOGIE:** TruSeq NE

## La Théorie



# Les Ressources

## 1 Génome du chêne

Plomion et al. 2018. *Nature Plants*

LETTERS

<https://doi.org/10.1038/s41477-018-0172-3>

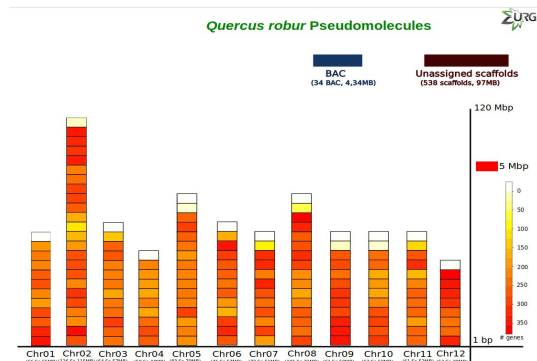


### Oak genome reveals facets of long lifespan

Christophe Plomion<sup>1,24\*</sup>, Jean-Marc Aury<sup>2,24</sup>, Joëlle Amselem<sup>3,24</sup>, Thibault Leroy<sup>1</sup>, Florent Murat<sup>4</sup>, Sébastien Duplessis<sup>5</sup>, Sébastien Faye<sup>2</sup>, Nicolas Francillonne<sup>3</sup>, Karine Labadie<sup>2</sup>, Grégoire Le Provost<sup>1</sup>, Isabelle Lesur<sup>1,6</sup>, Jérôme Bartholomé<sup>1</sup>, Patricia Favier-Rampant<sup>7</sup>

→ 44K gènes dont 25K positionnés sur les 12 pseudomolécules

[https://urgi.versailles.inra.fr/WebApollo\\_oak\\_PM1N/PseudoMolecule.html](https://urgi.versailles.inra.fr/WebApollo_oak_PM1N/PseudoMolecule.html)



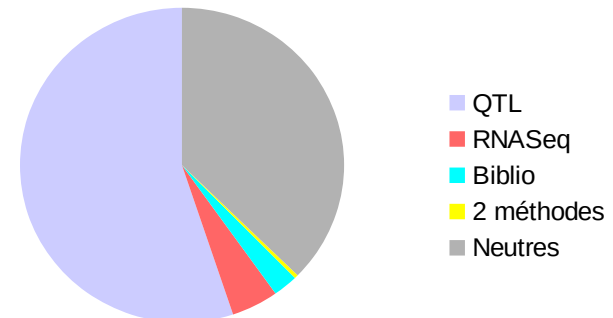
## 2 SNPs

→ chêne sessile en test de provenances  
→ gradient altitudinal et latitudinal en Europe



37,062,111 SNPs identifiés sur 18 pools de sessiles

## 3 1,186 Gènes candidats



## Les SNPs

### Choix des SNPs géniques:



■ exon  
□ intron

■ SNP sélectionné  
■ SNP non sélectionné

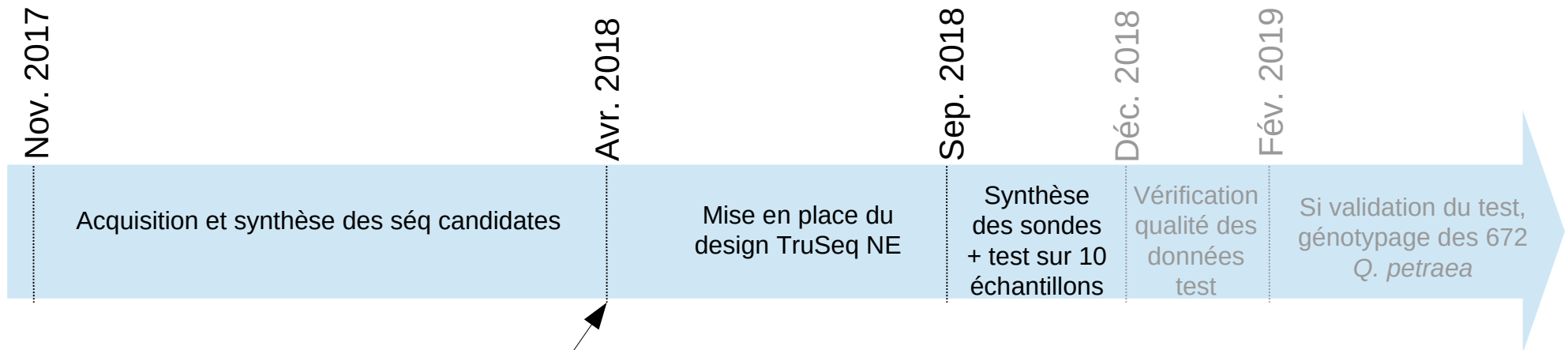
→ **89,007 SNPs dans 1,058 gènes**

### Choix des SNPs intergéniques:

- localisés dans des zones du génome a priori non sélectionnées
- 1 SNPs par scaffold
- scaffolds ancrés sur le génome (total = 870)
- à plus de 50bp de l'extrémité du scaffold

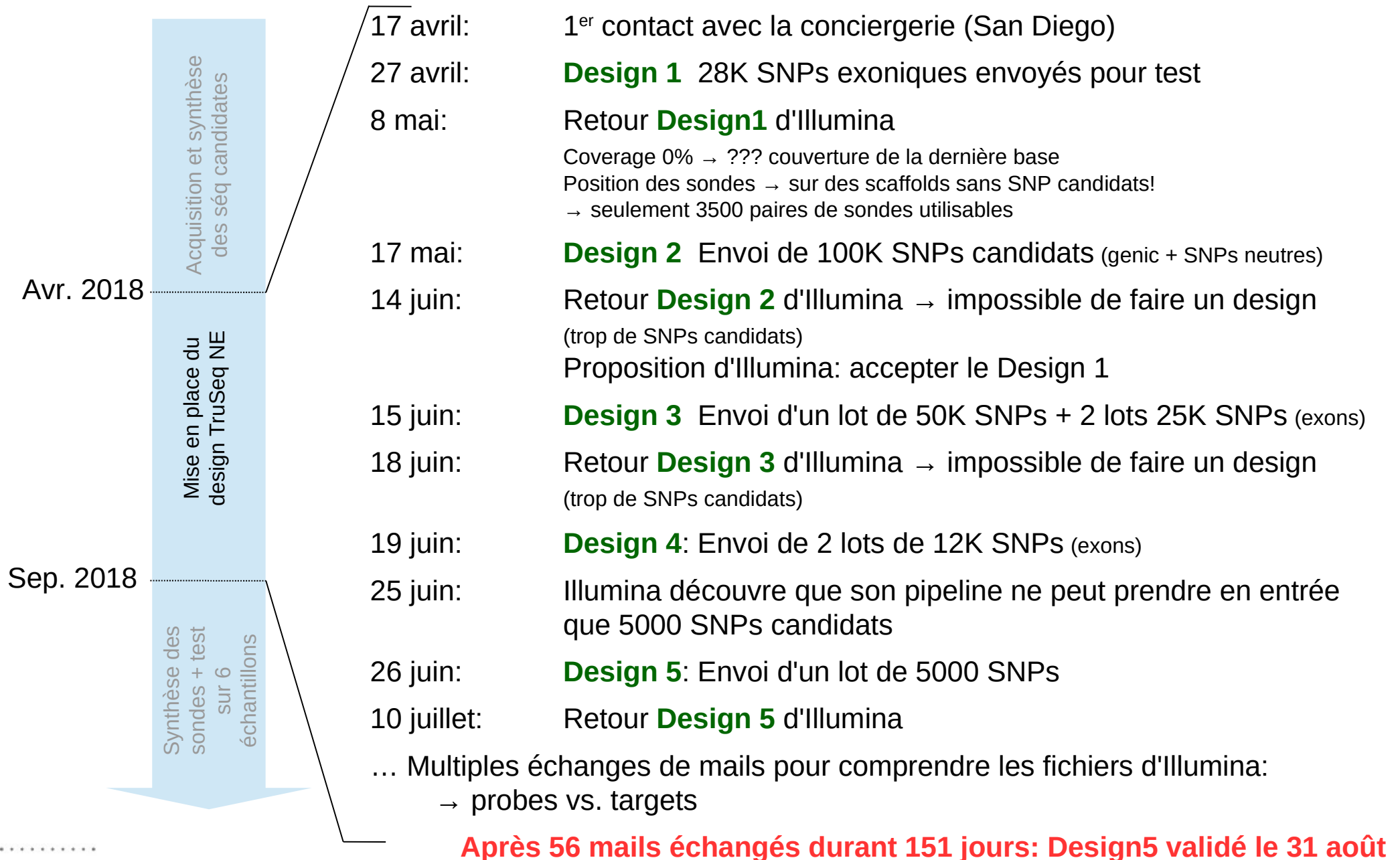
→ **709 SNPs dans 709 scaffolds**

# Le calendrier



*“Organisms that express high polymorphism can be challenging, unfortunately we don’t know how challenging until we start the design and perform functional testing using Oak DNA.”*

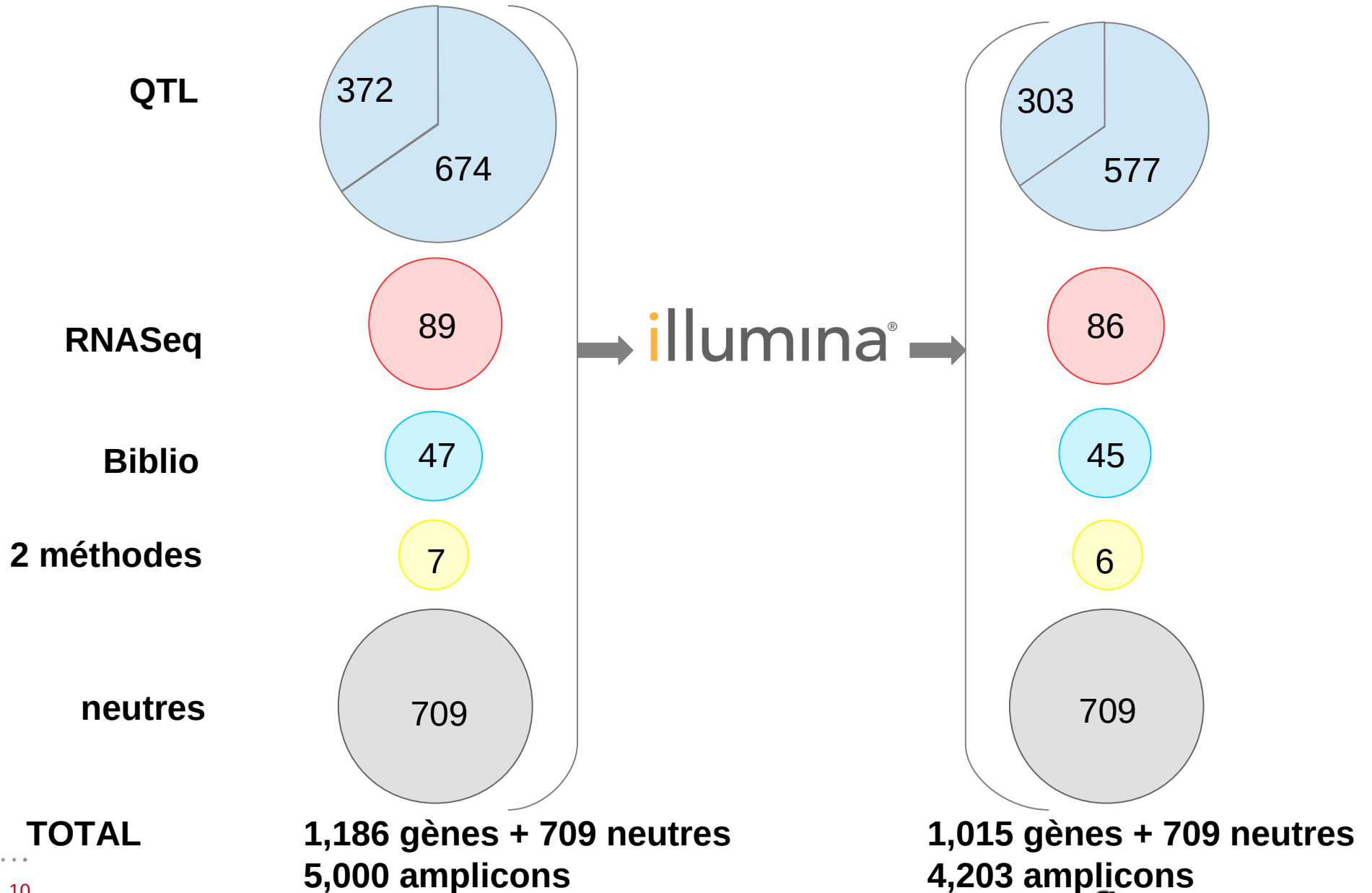
*Ivan Godínez, Illumina concierge*





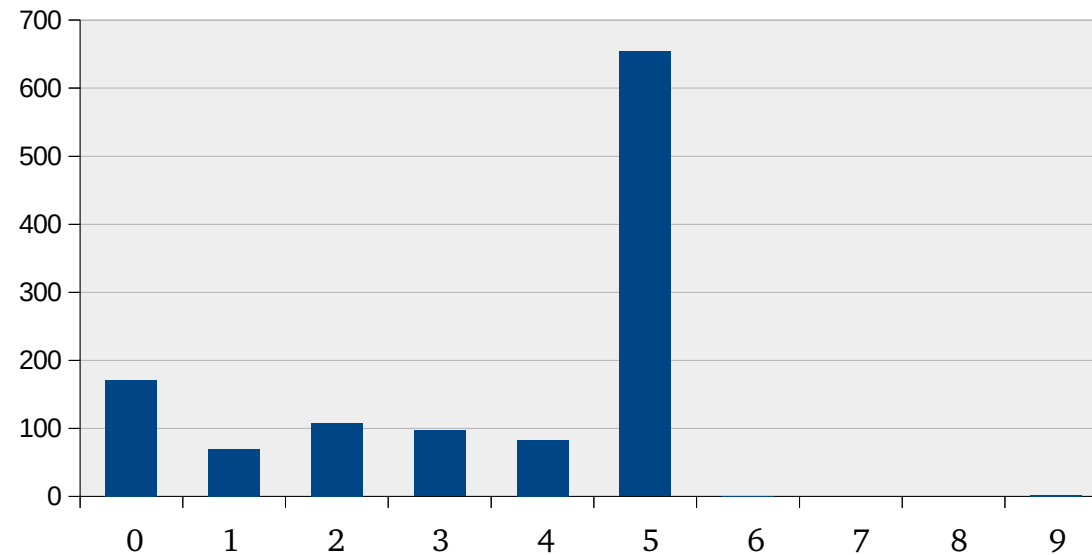
Séquences candidates sélectionnées

Design proposé par Illumina



## Distribution des sondes par gène

Nombre de gènes



Nombre de sondes

### Les +

- Le coût
- réactivité de la conciergerie Illumina
- possibilité de tester les sondes sur 10 individus (5 x 2) avant synthèse à grande échelle (échelle de qualité d'ADN + duplicats)

### Les -

#### Technologie:

- génotypage  $\neq$  capture: succès conditionné à la détection préalable de SNPs
- impossible d'atteindre les 5,000 amplicons
- pas de prise en compte des SNPs environnants pour le design des sondes

#### Echanges avec Illumina:

- nécessité d'avoir une personne qui vérifie le design
- langage différent (target  $\neq$  probe)
- absence de norme dans le format de fichiers
- difficulté d'obtenir des réponses (pourquoi des sondes sont positionnées sur des scaffolds sans SNP candidat ??)