



RÉPUBLIQUE FRANÇAISE

Liberté
Égalité
Fraternité



Isabelle Le Clainche, Aurélie Berard, Romain Mesnil, Damien Hinsinger, Patricia Faivre Rampant

INRAE unité US1279 - EPGV, Evry, France

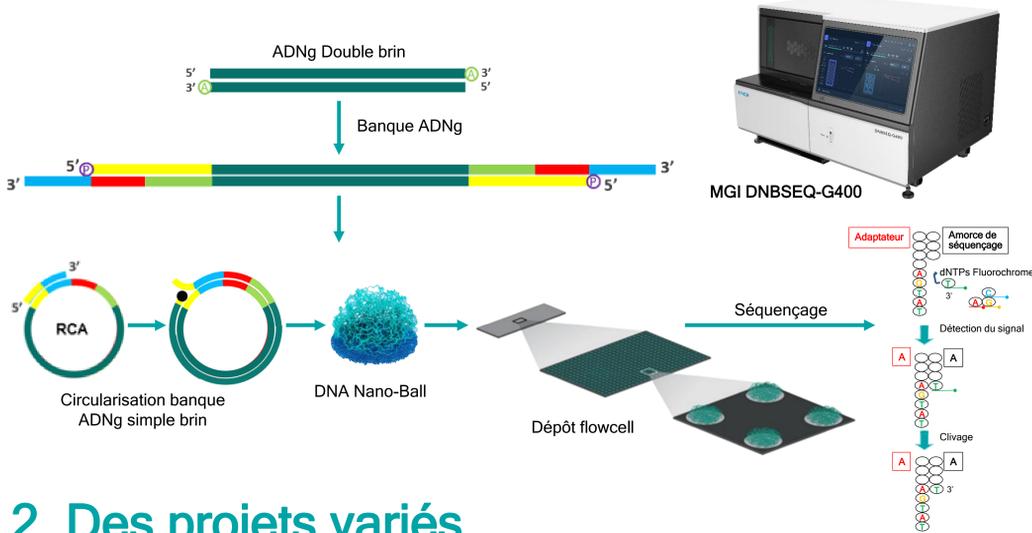
INRAE

EPGV
Etude du Polymorphisme
des Génomes Végétaux

INRAE
GENOMICS

> Un an de MGI à l'EPGV

1. Technologie MGI



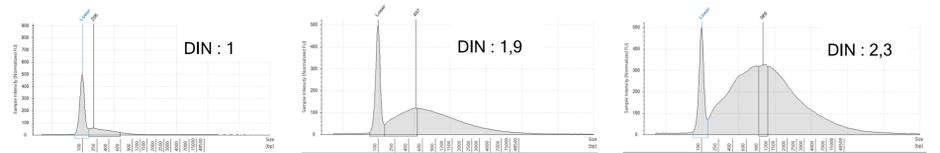
4. Le projet PoolKER



Séquencage en pool d'ADN de truite (*Salmo trutta*) pour l'étude de la diversification génomique lors de la colonisation des îles Kerguelen.

- Taille du génome : $\approx 2,3$ Gb
- Couverture visée de 50X
- Echantillonnage de 1973 à 2019, conservation dans du papier à température ambiante
- Extractions ADN individuelles à partir des tissus dermiques et de bases d'écaillés
- 14 pools équimolaires contenant 15 à 26 individus (réalisés par EcoBiop)
- ADN dégradé (90% des fragments < 600pb, voir exemples profils TapeStation ci-dessous)
- 10 échantillons >200ng; 4 échantillons <200ng

Profils TapeStation d'extraction d'ADNg de 3 pools :



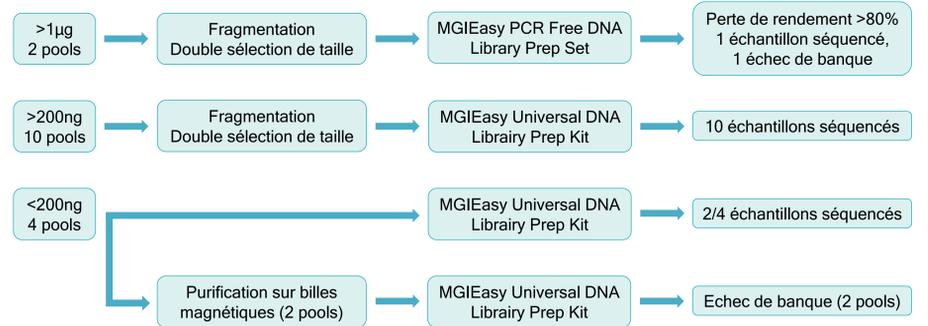
2. Des projets variés

340 échantillons végétaux, animaux et fongiques séquencés (construction manuelle des banques).

Echantillon	Taille du génome (Mb)	Nombre d'échantillons	Nombre de runs par kit de séquencage	Nombre de pistes
Peuplier	485	3	HotMPS 8,5	1
Mais	2 400	3		2
Tomate	950	3		2
Piment d'Espelette	3 500	10		10
<i>Lyallia kerguelensis</i>	1 600	1		1
Taureau domestique	3 000	3		2
<i>Phytophthora capsici</i>	110	48		2
Melon	380	253		18
Vigne	480	10		3
Truite	2 300	12		5
Total	/	346	13,5	54



Workflow projet PoolKER



3. Un rendement supérieur à l'attendu

Echantillon	Taille du génome (Mb)	Couverture visée (X)	Couverture obtenue $\geq Q30$ (X)	Rapport couverture obtenue/visée (%)
Peuplier	480	55	84	+53%
Mais	2 400	30	37	+23%
Tomate	950	20	10,6	-47%*
		100	68,6	-31%*
Piment d'Espelette	3 500	30	36,1	+20%
<i>Lyallia kerguelensis</i>	1 600	30	34,8	+16%
Taureau domestique	3 000	30	33	+10%
<i>Phytophthora capsici</i>	110	30	43,6	+45%
Melon	380	20	28,9	+45%
Vigne	480	13	20,0	+54%
Truite	2 300	50	55,7	+11%
/	/	/	Moy. rapport couverture obtenue/visée (%)	+18%

* Problème au niveau de la flowcell



5. Quelques grains de sable...

- Support de flowcell défectueux => arrêt en cours séquencage
- Valve et système d'évacuation des déchets remplacés
- Panne du disque dur
- Changement de version de kit, HotMPS vers StandardMPS, logiciel upgradé
- Chute de la qualité du read 2 relevée sur certains runs StandardMPS -> nouveau kit à venir

6. Conclusions

Les points négatifs :

- Préparation des banques longues et manuelles mais possibilité d'automatisation
- Choix complexe des index en amont pour équilibrer les bases
- Ajout de réactifs dans la cassette de séquencage
- Lavages non gérés automatiquement par le séquenceur DNBSEQ-G400
- Interventions régulières nécessitant l'immobilisation de la machine

Les points positifs :

- Coût du séquencage plus faible qu'Illumina®, optimal pour des petits projets
- Rendement supérieur à l'attendu (en moyenne +18%), en 1 an près de 5 Tb $\geq Q30$ produites, 90,5% en moyenne de bases $\geq Q30$
- Possibilité de conversion de banques Illumina® vers MGI lors du séquencage (à tester)

La technologie MGI est en progression constante et permet d'obtenir des données de très bonne qualité, malgré quelques difficultés opérationnelles.

Center
Île-de-France - Versailles-Saclay

Remerciements



INRAE - Unité EPGV US1279
CEA - Institut de biologie François Jacob
Site d'Evry - Bat G1
2 rue Gaston Crémieux
91057 - Evry Cedex
France
support-epgv@inrae.fr

<https://www6.versailles-grignon.inrae.fr/epgv>