


---

## Recommandation de l'utilisation du Kit Illumina DNA PCR-Free Prep avec échantillon non tumoral

---

### Table des matières

1	OBJET .....	2
2	INTERETS POUR LES PLATES-FORMES DU PLAN FRANCE MEDECINE GENOMIQUE.....	2
3	CHAMP D'APPLICATION DE LA RECOMMANDATION .....	2
4	PRESENTATION DU KIT.....	2
4.1	Principe .....	2
4.2	Avantages décrits par le fournisseur .....	2
4.3	Protocoles selon quantité de départ.....	3
5	PLAN DE TESTS.....	3
5.1	Echantillons.....	3
5.2	Expériences .....	3
6	RESULTATS.....	4
6.1	Qualité des Librairies .....	4
6.2	Qualité du Séquençage .....	4
6.3	Analyse des Variants .....	4
6.4	Analyse de la Similarité des Ensembles de Variants .....	6
7	LES POINTS DE VIGILANCES.....	6
7.1	Dosage et dépôts des librairies .....	6
7.2	Création du Pool .....	6
7.3	Index .....	7
7.4	Régions riches en GC.....	7
7.5	Capacité des Flow Cell.....	7
7.6	Nombre d'échantillons par Flow Cell et re-dépôts .....	7
8	VALIDATION PAR LES PLATES-FORMES.....	7
9	CONCLUSION GENERALE.....	7
10	DOCUMENTS.....	8
11	HISTORIQUE DES MODIFICATIONS .....	8
12	VALIDATION DU DOCUMENT.....	8

N°: 20221109_RT07 Version : 02 Date : 09/11/2022	Recommandation Technique PFMG 2025	 France Médecine Génomique 2025
Sujet	Illumina DNA PCR-Free Prep	
Page 2/8		

## 1 OBJET

Analyse du kit Illumina DNA PCR-Free Prep (Tag), kit de préparation de bibliothèques par fragmentation enzymatique et sans amplification à partir de faible quantité d'ADN de départ (25 - 300 ng) pour du séquençage de génome entier.

Cette analyse est réalisée par comparaison avec le kit TruSeq DNA PCR-Free (TruSeq) qui est principalement utilisé dans le plan aujourd'hui et des jeux de données externes.

## 2 INTERETS POUR LES PLATES-FORMES DU PLAN FRANCE MEDICINE GENOMIQUE

Ce kit permettrait :

- **D'abaisser la quantité de matrice** de départ nécessaire pour réaliser des bibliothèques, permettant de séquencer les échantillons avec peu d'ADN sans avoir recours à l'amplification par PCR,
- Un **gain de temps**, lors de la préparation de bibliothèques et **seulement un dosage du pool** final des bibliothèques dans le cas d'une quantité supérieure ou égale à 300 ng,
- Un **gain économique**, l'étape de fragmentation ne demandant plus d'appareil spécifique (ni maintenance).

## 3 CHAMP D'APPLICATION DE LA RECOMMANDATION

- Cette recommandation s'applique à la création de bibliothèques sans amplification pour des échantillons d'ADN constitutionnel permettant le séquençage de leur génome entier (WGS) sur plateforme Illumina.
- Cette recommandation ne s'applique ni aux ADN extraits à partir de tissu fixé (ex : FFPE), ni aux échantillons tumoraux.

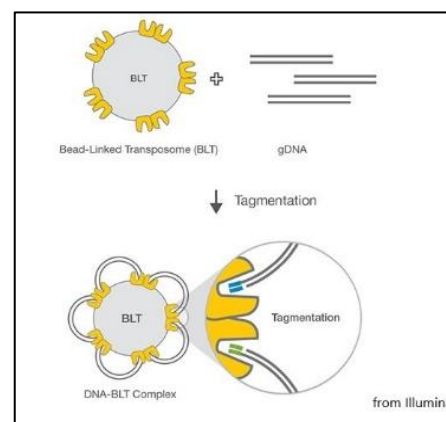
## 4 PRESENTATION DU KIT

### 4.1 Principe

Utilisation de transposases accrochées sur des billes magnétiques permettant une coupure des brins d'ADN.

Ces transposases coupent de façon aléatoire l'ADN et ajoutent une partie d'adaptateur sur les extrémités finales. La répartition des transposases sur les billes permet d'obtenir des coupures de tailles homogènes.


La quantité de billes par réaction est fixe, les billes sont saturées dès que la quantité d'ADN dépasse les 300 ng.



### 4.2 Avantages décrits par le fournisseur

Selon Illumina, les avantages de ce kit par rapport au TruSeq DNA PCR-Free et autres kits concurrents sont :

- Une quantité de matrice de départ plus faible de 25 à 300 ng au lieu de 1000 ng (TruSeq), avec toutefois des modes opératoires différents selon la quantité.
- Un gain de temps de manipulation, par réduction des étapes de purification (de 3 à la place de 4).
- Une normalisation simplifiée du pool de bibliothèques dans le cas d'une quantité supérieur à 300 ng.
- Un gain de couverture dans les régions riches en CG.

N°: 20221109_RT07 Version : 02 Date : 09/11/2022	Recommandation Technique PFMG 2025	 France Médecine Génomique 2025
Sujet	Illumina DNA PCR-Free Prep	
Page 3/8		

### 4.3 Protocoles selon quantité de départ

Quantité départ (ng)	Normalisation ADN	Protocole	Normalisation librairie	Dépôts sur Novaseq
≥ 300	Non	Standard	Dosage du pool final	Standard
100-299	Oui	Standard	Dosage de chaque librairie	Standard
25-99	Oui	Low input	Dosage de chaque librairie	Mode XP

## 5 PLAN DE TESTS

### 5.1 Echantillons

Les expériences ont été réalisées sur un panel de 7 échantillons standards, dont les SNV (Single Nucleotide Variant) sont publiés et référencés pour cinq d'entre eux.

Description		Nom	ID coriell	Filiation	Reference	Provenance
CEPH/UTAH PEDIGREE 1463	TRIO	CEPH146302	NA12878	Fille	GIAB	CEPH
		CEPH146315	NA12891	Père		CEPH
		CEPH146316	NA12892	Mère		CEPH
PERSONAL GENOME PROJECT	TRIO	Ashkénaze	NA24385	Fils	GIAB	Coriell
		Ashkénaze	NA24149	Père	GIAB	Coriell
		Ashkénaze	NA24143	Mère	GIAB	Coriell
		Han (Chinois)	NA24631	Fils	GIAB	Coriell

Ces ADN ont été achetés auprès du Coriell Institut et du CEPH.

### 5.2 Expériences

Les tests ont porté sur la variation de la quantité de matrice (25, 100, 300, 500, 1000 ng). Les librairies ont été faites en duplicat pour le trio CEPH.


Expériences	Kit	Nb d'échantillons	Quantité en ng par échantillon	Dosage des librairies	Type de FC	Profondeur cible
TruSeq 1000	TruSeq DNA PCR-Free	7 + 3 <sup>1</sup>	1000	Oui	S2 300 cycles	30 X
Tag 500	Illumina DNA PCR-Free Prep	7 + 3 <sup>1</sup>	500	Non	S2 300 cycles	
Tag 300		7 + 3 <sup>1</sup>	300	Oui <sup>2</sup>		
Tag 100		7	100	Oui		
Tag 25		7	25	Oui	S2 300 cycles en mode XP	

- <sup>1</sup> Librairies faites en duplicat puis déposées.

- <sup>2</sup> Déviation par rapport au protocole car variation sur les concentrations de librairies obtenues.

Les librairies ont été déposées uniquement sur un Novaseq 6000 afin d'éviter toute dérive liée au séquenceur, avec des Flow Cell (FC) de type S2. Le CRefIX a fait le choix de déposer en FC S2 pour des raisons budgétaires et de temps.

La profondeur choisie est de 30 X comparable à celle couramment utilisé par le plan FMG 2025 pour les mêmes échantillons. En parallèle, ces échantillons ont été séquencés sur le même type de FC à partir de librairies créées en TruSeq DNA PCR-Free.

N°: 20221109_RT07 Version : 02 Date : 09/11/2022	Recommandation Technique PFMG 2025	 France Médecine Génomique 2025
Sujet	Illumina DNA PCR-Free Prep	
Page 4/8		

## 6 RESULTATS

Le CRefIX a analysé :

- La qualité et la quantité de librairies obtenues,
- Le nombre de séquences générées et leur qualité,
- Les variants obtenus qui ont été comparés avec les séquences issues des bases de données (ref dans GIAB - Régions dites "à haute confiance") et celles issues du séquençage en TruSeq DNA PCR-Free.

### 6.1 Qualité des Librairies

Critères	Observés		Attendus (valeurs illumina)	Résultats
Gamme de concentration (nM)	6-10 7 -10 2- 4 0,75 – 1,9	500 ng 300 ng 100 ng 25 ng	7 nM (300 ng)*	✓
Gamme de taille d'insert (pb base)	450 +/- 10 480 +/- 20 400 +/- 14 400 +/- 10	500 ng 300 ng 100 ng 25 ng	450 +/- 70 *	✓
Profil	Homogène		Homogène	✓

\* Valeurs données par Illumina pour un input de 300 ng.

### 6.2 Qualité du Séquençage

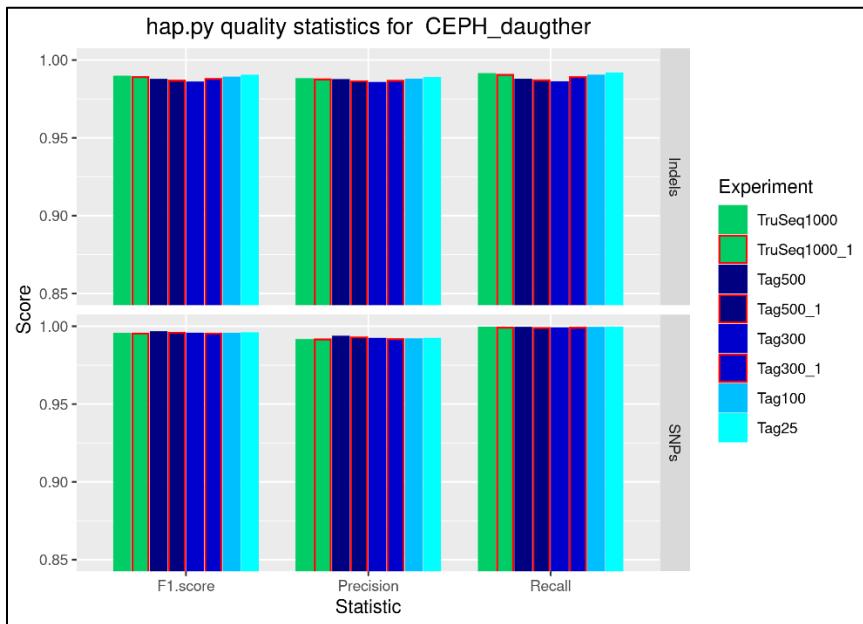
Critères	TruSeq / CNRGH *	Illumina DNA PCR-Free Prep	Attendus (valeurs illumina)	Résultats
Phasing, prephasing	< 0,1	< 0,1	< 0,6	✓
Q30	> 90 %	90 %	> 75 %	✓
Rendement total (Gb)	1350,5	1343	1333	✓
Taux de duplicat <sup>1</sup>	< 10 %	12 – 22 %	-	✘
Couverture	>99,2%	>99,2%	-	✓
Profondeur	~30X homogène	~30X homogène	-	✓
Profondeur GC riche <sup>2</sup>	~30X homogène	~20X homogène	↗	Perte de 10X

\* Valeurs moyennes obtenues par le Centre National de Recherche en Génomique Humaine (CNRGH).

- <sup>1</sup> Taux de duplicats plus important, constaté en Illumina DNA PCR-Free Prep, cette perte selon Illumina peut-être compensée par un chargement optimisé des FC.
- <sup>2</sup> Perte de 10 x de profondeur dans les régions de plus de 85 % de GC (représentant 0.02 % du génome).

### 6.3 Analyse des Variants

Présentation des résultats des SNV pour l'individu CEPH 12878 sur l'ensemble des expériences. L'ensemble des séquences a été « downsampled » à 30 X pour homogénéiser les analyses.



Graphique de la précision (taux de faux positifs, « Precision »), de la sensibilité (taux de faux négatifs, « Recall ») et du F1 score pour l'individu CEPH selon les expériences.

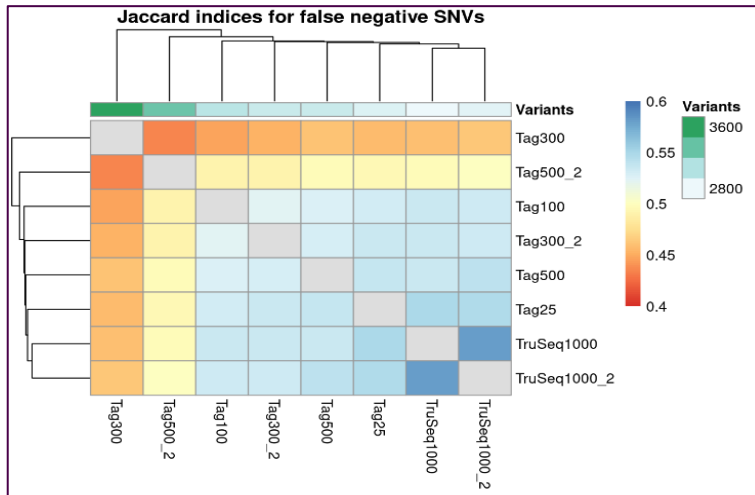
Critères		TruSeq DNA PCR-Free	Illumina DNA PCR-Free Prep	Résultats
Nombre de variants		3,5M	3,5M	✓
Type de variants		86% SNV	86% SNV	✓
Qualité des SNVs	Precision	99,1%	99,1 - 99,3%	✓
	Recall	99,9%	99,8 - 99,9%	✓
	F1 Score	99,5%	99,5 - 99,6%	✓
Qualité des Indels	Precision	98,7 %	98,5 - 98,8 %	✓
	Recall	99 %	98,6 - 99,1 %	✓
	F1 Score	98,9 %	98,5 - 99 %	✓
Distribution des variants selon leur fréquence allélique		TP* : pic à 0,5 et 1 FP* : pic à 0,1	idem	✓
Qualité des SNVs des régions GC riches	Precision	70 - 76,5%	71,7 - 80,5%	✓
	Recall	99,6%	98,3 - 99,6%	≈
	F1 Score	82,2 - 86,5%	83,3 - 88,9%	✓

Tableau de synthèse de la qualité des variants par type.

\* TP = True Positive, FP = False Positive

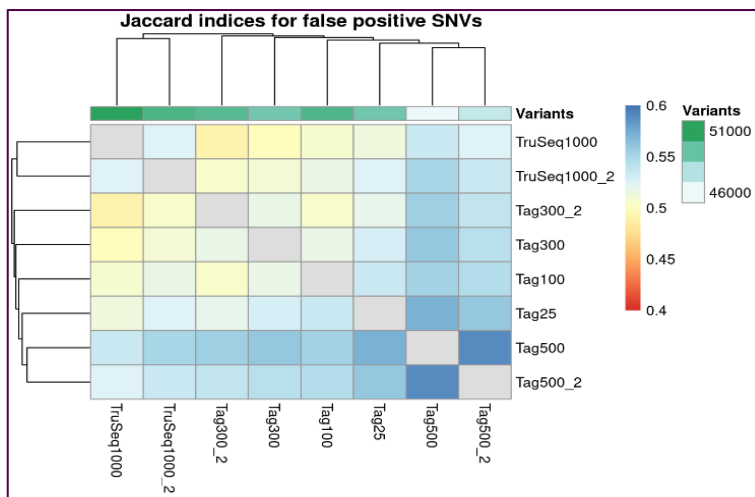
## 6.4 Analyse de la Similarité des Ensembles de Variants

Recherche de biais par mesure de similarité (Indice de Jaccard) pour l'individu CEPH 12878.



Similarité des Ensembles de SNV Faux Négatifs.

Les faux négatifs des résultats TruSeq DNA PCR-Free sont légèrement plus proches entre eux. La variation pour l'Illumina DNA PCR-Free Prep ne montre pas de faux négatifs spécifiques à ce kit.



Similarité des Ensembles de SNV Faux Positifs.

Les faux positifs (artéfacts) ne montrent pas de cluster net séparant Illumina DNA PCR-free Prep versus TruSeq DNA PCR-Free.

L'analyse ne montre pas de biais lié au kit.

## 7 LES POINTS DE VIGILANCES

### 7.1 Dosage et dépôts des bibliothèques

Le calcul de la quantité de bibliothèques générée est une étape cruciale. Le CRefIX attire l'attention des futurs utilisateurs sur cette étape. Le CRefIX recommande de doser par une méthode de qPCR.


La quantité optimale de pool à déposer sur les FC sera à ajuster en fonction des plates-formes et des appareils utilisés.

Le CRefIX attire l'attention sur le fait que les bibliothèques créées sont en simple brin, et qu'il n'est donc pas possible de vérifier le profil de migration. D'ailleurs, Illumina ne préconise pas de migration pour vérifier les tailles.

Les FC doivent être de la version 1.5. Dans le cas de l'utilisation d'une version antérieure, il faudra rajouter un custom primer (VP 10) à la cassette de séquençage.

### 7.2 Création du Pool

Dans le cas d'une quantité de départ supérieure à 300 ng, Illumina considère que le dosage du pool final est suffisant. Le CRefIX recommande de doser chaque bibliothèque afin d'éviter tout déséquilibre de pool engendrant des profondeurs différentes pour les échantillons.

N°: 20221109_RT07 Version : 02 Date : 09/11/2022	Recommandation Technique PFMG 2025	 France Médecine Génomique 2025
Sujet	Illumina DNA PCR-Free Prep	
Page 7/8		

### 7.3 Index

Selon l'application note d'Illumina (*Balancing sample coverage for whole-genome sequencing*), il y a un risque de perte ou d'augmentation du rendement du séquençage selon les index utilisés.

N'ayant testé qu'un nombre limité d'index, le CRefIX n'a pas vu d'impact sur ses expériences.

### 7.4 Régions riches en GC

On note une perte de 10 X en profondeur sur les régions à plus de 85 % de GC (qui représentent moins de 0.02% du génome).

### 7.5 Capacité des Flow Cell

Le taux de duplicat étant plus élevé qu'un TruSeq DNA PCR-Free, la productivité des FC en séquences utiles s'en trouve réduite. Dans le cas d'une S2, le nombre d'échantillon ayant pour cible une profondeur de 30 X, passera de 8 à 7.

### 7.6 Nombre d'échantillons par Flow Cell et re-dépôts

Comme les concentrations et les volumes obtenus des bibliothèques sont plus faibles par rapport au gold standard « TruSeq DNA PCR-Free », le CRefIX attire l'attention sur le nombre d'échantillons et la non possibilité de déposer de nouveau en fonction du type de FC utilisée et de la profondeur désirée. D'ailleurs les volumes et concentrations des bibliothèques obtenues lors du test à 25 ng, ne permettent pas pour certaines bibliothèques un re-dépôt même en mode XP.

## 8 VALIDATION PAR LES PLATES-FORMES

Le CRefIX n'étant pas un organisme de validation, les tests de répétabilité, de robustesse, de limite de détection (non applicable) n'ont pas été faits.

Les plates-formes désirant utiliser ce kit devront refaire une validation de méthode.

Le CRefIX n'a pas réalisé de test en utilisant une plate-forme robotique. Au vu de la variabilité du parc d'instruments de pipetage, le CRefIX n'a pas la possibilité de tester tout le parc. Mais l'analyse des différentes étapes du protocole ne semble pas poser de problème majeur pour une automatisation sur une plate-forme robotique.


## 9 CONCLUSION GENERALE

L'ensemble des critères sont dans les spécifications d'Illumina, pour chaque expérience.

L'analyse des variants de l'ensemble des échantillons ne montre pas de différence par rapport à un TruSeq DNA PCR-Free.

Critères	Truseq DNA PCR-Free	Illumina DNA PCR-Free Prep	Commentaires
Quantité	☹️ --	😊 +++	1/3 en moins.
Temps	☹️ --	😊 +++	1h30 au lieu de 5h00.
Automatisable	😊	😊	Cf Illumina.
Normalisation	😐	☹️	Reste à valider.
Couverture	😐	😐	Couverture identique voire plus homogène. Mais perte de 10x de profondeur dans les régions riches en GC (0.02 % du génome)
Variants	😊	😊	Equivalents

**Au vu des avantages liés à une plus faible quantité de départ et une réduction du temps, le CRefIX considère que ce kit peut être utilisé dans le cas d'échantillons non tumoraux sous réserve des points de vigilance.**

N°: 20221109_RT07 Version : 02 Date : 09/11/2022	Recommandation Technique PFMG 2025	 France Médecine Génomique 2025
Sujet	Illumina DNA PCR-Free Prep	
Page <b>8/8</b>		

## 10 DOCUMENTS

- Illumina DNA PCR-Free Reference Guide :  
Comprehensive information on Illumina DNA PCR-Free Library Prep, including a detailed protocol. Ref Document # 1000000086922 v03 (February 2021)
- Illumina DNA PCR-Free Checklists :
  - Illumina DNA PCR-Free Library Prep Standard Input Checklist Ref Document # 1000000086923 (Jul 21, 2020)
  - Illumina DNA PCR-Free Library Prep Low Input Checklist Ref Document # 1000000130333 (Jul 21, 2020)
  - Illumina DNA PCR-Free Library Prep Hybex Protocol Checklist Ref Document # 1000000130364 (Jul 21, 2020)
- Bruinsma S, et al. Bead-linked transposomes enable a normalization-free workflow for NGS library preparation. BMC Genomics. 2018 Oct 1;19(1):722.

## 11 HISTORIQUE DES MODIFICATIONS

Version	Date	Nature de la modification
V02	09/11/2022	Approbation
V01	01/11/2022	Revue et correction
V00	01/06/2022	Création

## 12 VALIDATION DU DOCUMENT

RÉDACTEURS	RELECTEURS
Nom : Edouard Turlotte, Jasmin Cévest	Nom : Alice Moussy, Margaux Gras, Marine Rouillon, Violette Turon, Alain Viari, Jean-François Deleuze
Fonction : Ingénieur, Bio-informaticien	Fonction : Ingénieur, Chercheur, Bio-informaticien
Date : 10/11/2022	Date : 10/11/2022
Visa :	Visa :