



Comparison of the PacBio 2.0 and 3.0 kits for full length mRNA sequencing in multiplex

Marine Sallaberry, Eden Darnige, Lucien Hoffbeck, Sandra Fourre, Mathis Pochon, Martine da Rocha, Sophie Mantelin, Cyril van Ghelder, Etienne Danchin, Karine Robbe-Sermesant, et al.

► To cite this version:

Marine Sallaberry, Eden Darnige, Lucien Hoffbeck, Sandra Fourre, Mathis Pochon, et al.. Comparison of the PacBio 2.0 and 3.0 kits for full length mRNA sequencing in multiplex. Assemblée Générale France Génomique, Jun 2023, Paris, France. hal-04150252

HAL Id: hal-04150252

<https://hal.science/hal-04150252>

Submitted on 4 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Comparison of the PacBio 2.0 and 3.0 kits for full-length mRNA sequencing in multiplex

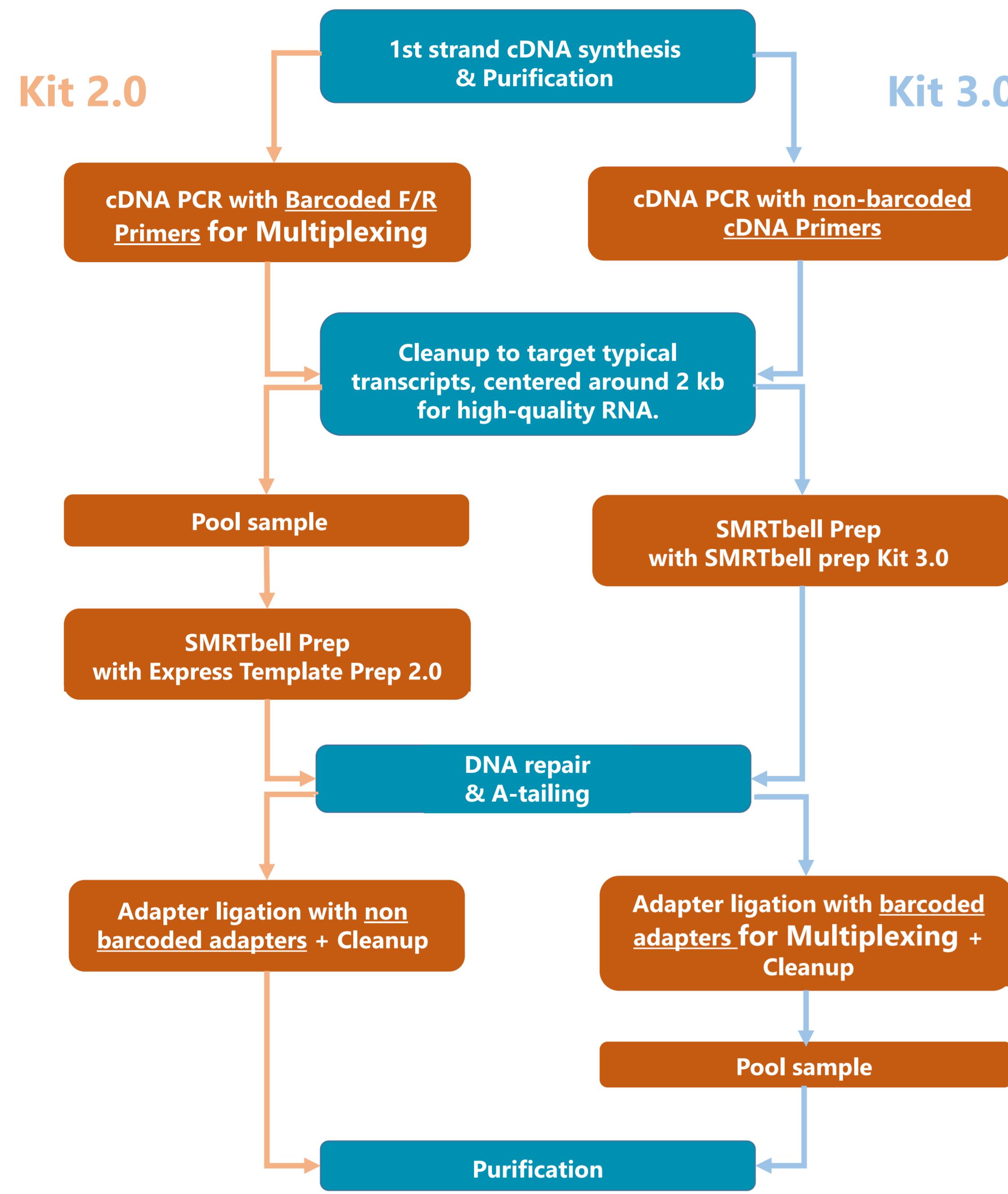
Marine SALLABERRY¹, Eden DARNIGE¹, Lucien HOFFBECK¹, Sandra FOURRE¹, Mathis POCHON², Martine DA ROCHA², Sophie MANTELIN², Cyril VAN GHELDER², Etienne DANCHIN², Karine ROBSE-SERMESANT², Mohammed Zouine³, Carole IAMPIETRO¹, Céline VANDECASSELE¹, and Cécile DONNADIEU¹

¹ GeT-PlaGe, UMR 1426, INRAE, Genotoul, Castanet-Tolosan ; ² Institut Sophia Agrobiotech, UMR1355 INRAE-Université Côte d'Azur-CNRS; ³ Génomique et biotechnologie des fruits, UMR 990, INRAE

Abstract

Our aim is to compare 2 Iso-Seq protocols carried out at the GeT-PlaGe core facility following an evolution of the kits offered by PacBio. For both protocols, we started with 4 mRNAs. The main difference is the timing of the addition of barcodes to the samples. The protocol with the Express Template Prep 2.0 (Kit 2.0) proposes early sample identification, whereas the protocol with the SMRTbell prep kit 3.0 (Kit 3.0) only includes the barcodes at the end of library preparation. Each protocol has its advantages and disadvantages, which we will describe in detail below. For this analysis, we relied on 2 scientific projects, ISOSOL and CHIMIONEM, in collaboration with the BMMF Genomics and Fruit Biotechnology laboratory and the Institut Sophia Agrobiotech (ISA) laboratory respectively. We must take into consideration that we are comparing different samples from different projects; the comparison is for informative purposes only. The ISOSOL project contains 4 tobacco samples, and the CHIMIONEM project contains 4 nematode samples.

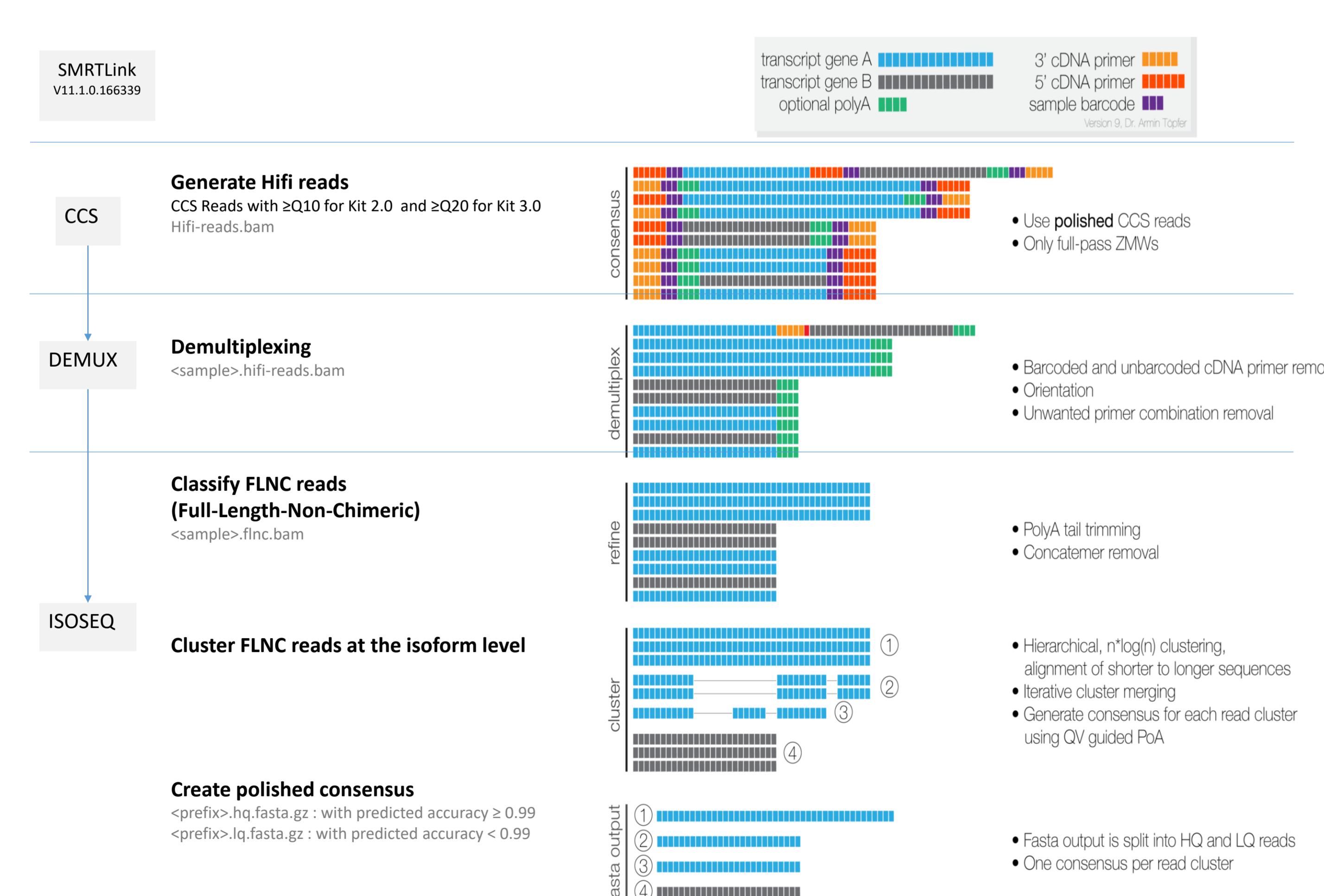
PacBio's kit 2.0 and 3.0 protocols



Kit 2.0 is represented by the orange path. Iso-Seq library preparation using the SMRTbell® Express Template Prep-Kit 2.0 was used by GeT-PlaGe in 2021. After reverse transcription, it consists of amplifying the fragments of obtained cDNA (including 15 PCR cycles) including barcoded cDNA primers which make it possible to make a pool at the beginning of the library, thus allowing to build only one pool gathering all of the samples. We tested this protocol on 4 RNAs with 300ng of input. Purification was done with PronEx beads.

Kit 3.0 is represented by the blue path. Iso-Seq library preparation using the SMRTbell® prep Kit 3.0 was used by GeT-PlaGe in 2023. After reverse transcription, it consists of amplifying the fragments of obtained cDNA (including 15 PCR cycles), including non-barcoded cDNA primers, thus allowing to build each sample with SMRTbell Barcoded Adapter Plate 3.0 at the end of the library preparation. We tested this protocol on 4 RNAs with 300ng of input. Purification was done with CleanUp beads. The option of multiplexing with barcodes at the beginning of the library is also possible with the Kit 3.0, but we have not yet tested this.

ISO-SEQ analysis workflow details

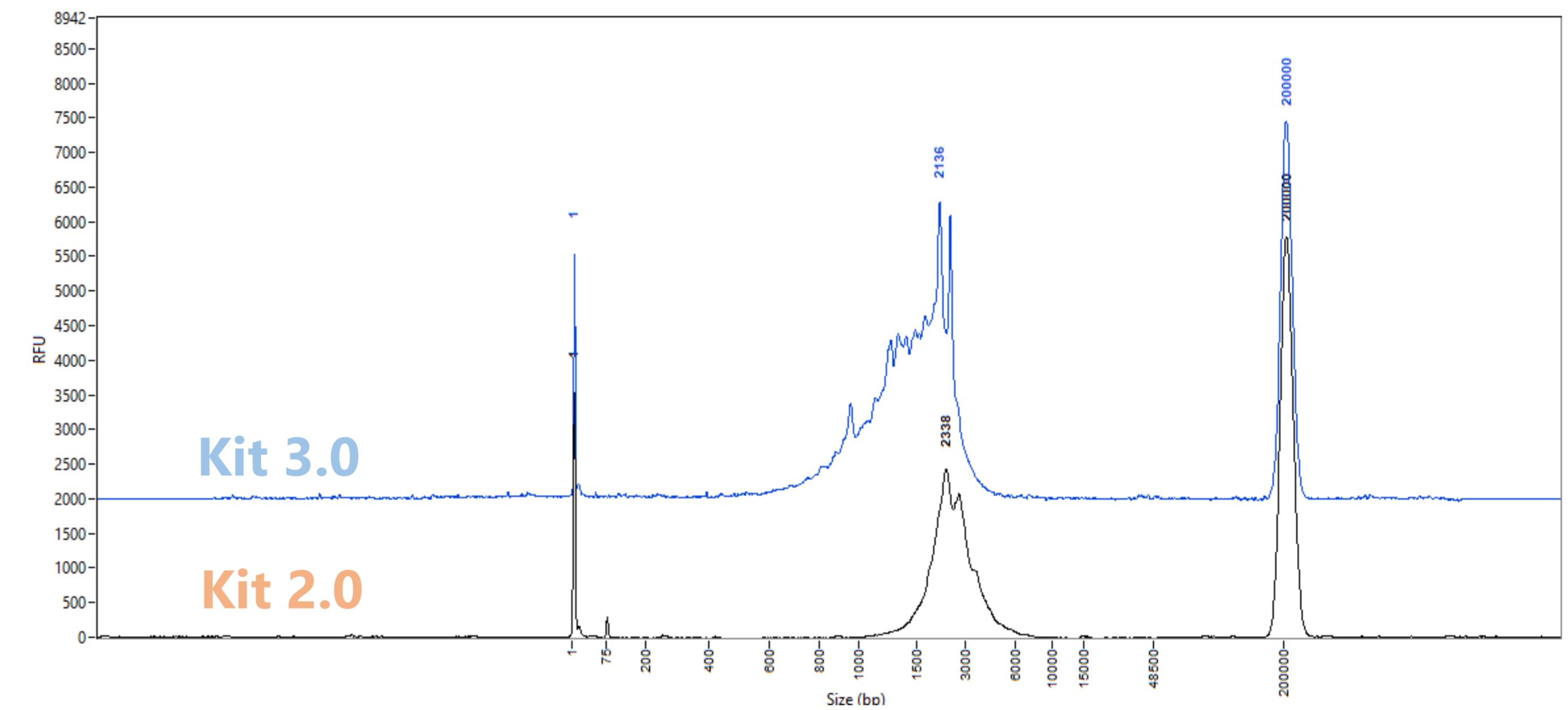


This workflow is valid for both protocols.

Conclusion

In the end, we obtained better results with Kit 3.0 in terms of data quantity, pool balance and number of unclassified reads. However, these results may be called into question should a project involving more samples be considered. Indeed, multiplying the number of libraries in the second protocol decreases the time spent on library construction, even if the Kit 2.0 reduces the number of tubes required. Additionally, Kit 3.0 increases the cost, which is not negligible. It is important to bear in mind that these are not the same samples used for both protocols, which adds a bias to the results that could be sample-dependent. This trend has yet to be verified with other projects.

Profiles of the libraries obtained



Kit 2.0. The profile is slightly flared on both sides of the central peak, we also notice the presence of a peak at 70 bp. We chose to select typical transcripts, centered around 2 kb, for high-quality RNA. The average size of the pool is 2kb. The sizes of the 4 cDNAs before pooling were also around 2kb. The pool was made equimassically. Purifications and circularization of the libraries were done on the pool, so we do not have a view on the cDNAs independently before obtaining this profile. The profile was obtained using the Fragment analyser.

Kit 3.0. The profile is different from the one obtained with the kit 2.0, with a more thickened profile on the side of the small fragments. These differences may be due to differences in the samples or the method. We chose to select typical transcripts, centered around 2 kb, for high-quality RNA. The average size is 2kb which corresponds to the sizes obtained previously on each independent cDNA. The profile was obtained using the Fragment analyser. The pool was made equimassically. No purification or other steps were done on this pool to avoid the risk of unbalancing the pool without visibility.

Benefits / disadvantages between Kit 2.0 and Kit 3.0

	Cost ¹	Time ¹	Pool balance visibility	Barcode accessibility
Kit 2.0	V	X	X	X
Kit 3.0	X	V	V	V

Using the Kit 2.0 protocol, the cDNAs are pooled before construction of the classical library, allowing for circularization. Here, we use a quantity of reagent equivalent to a single reaction, whereas for the 3.0 protocol, each sample represents a reaction, therefore costs are increased. The price must also take into account the quantity of beads required for purification and consumables, which for Kit 3.0 are multiplied by the number of samples. In the same way, the time is slightly reduced for protocol 2.0 since there is only one tube to be treated as compared to a number of tubes equivalent to the number of samples for the 3.0 method. However, we note that Kit 2.0 has more steps than Kit 3.0, making the protocol longer to complete. From a laboratory perspective, an advantage of Kit 2.0 is the visibility of the samples independently until the last moment, which favors the obtaining of an optimally equilibrated pool. In terms of barcode accessibility, the 2.0 method requires custom designing based on the proposed PacBio kits, whereas the 3.0 method uses the 3.0 barcodes classically used with the major library preparation kits.

1 : Related to the number of samples to be taken, here we are referring to about 4 samples to be processed.

SMRT Link CCS/Demultiplex

Kit 2.0	Samples	# CCS Reads	% CCS	# FLNC reads
Tabacco 1	783 938	27,4%	782 728	
Tabacco 2	530 482	18,6%	523 068	
Tabacco 3	617 285	21,6%	616 344	
Tabacco 4	680 497	23,8%	679 110	
All samples	2 857 761	-	2 601 250	
Unclassified	245 559	8,6%	-	

Kit 3.0	Samples	# CCS Reads	% CCS	#FLNC reads
Nematode 1	841 008	21,44%	840 233	
Nematode 2	993 794	25,34%	992 500	
Nematode 3	829 082	21,14%	828 051	
Nematode 4	1 119 193	28,53%	1 117 535	
All samples	3 922 363	-	3 778 430	
Unclassified	139 286	3,55%	-	

We used two different batches of SMRTcells and sequencing reagents, but the quality of the runs remained comparable. FLNC : Full Length Non Chimeric.