eurofins | Genomics

APPLICATION NOTE

A new oligonucleotide synthesis process for high quality oligonucleotides with superior performance in NGS applications

A new oligonucleotide synthesis process for high quality oligonucleotides with superior performance in NGS applications

Max Leopold, Maximilian Himmelstoß, Anna Spanopoulou, Regina Dick, Samanta Cazorla, Johannes Kremser, Robert Galik, James Konynenbelt, Thomas Halbritter, Markus Schmitt

Abstract

During the last decade, Next Generation Sequencing (NGS) became established as the preferred technology for screening of genomic variants for diagnostics, clinical genomics and agrigenomics. To obtain optimal performance, high-quality oligonucleotides are needed which are usually purified by HPLC. The HPLC purification of longer oligonucleotides is tedious and prone to cross-contamination. In this application note, we describe a newly developed procedure to prepare unique dual index (UDI) primers. This fast workflow minimizes cross-contamination to levels below 0.002%. The synthesis technology and downstream process utilized produces high performance NGS oligonucleotides that will revolutionize the applicability and quantity of oligonucleotides in NGS.

Introduction

Next generation sequencing (NGS) is considered the most important technique for high throughput genetic testing. During the past decade, NGS gained popularity due to its revolutionary parallel sequencing capability to detect a variety of genomic alterations, e.g. mutations, gene amplifications, gene fusions and gene expression. The wide range of high precision applications made NGS one of the most preferred technologies in research, clinical environments, and breeding. Regardless of the NGS application or platform, high quality oligonucleotides are key for successful sequencing results.

Several types of NGS oligonucleotides are required for the different NGS applications, but Index Adapter Oligonucleotides stand out, because they must meet the highest quality requirements (**Table 1**). Index Adaptors contain specific identifier sequences (indices) that are added to the library fragments of the individual sample during library preparation. Differently indexed samples can be pooled and then simultaneously sequenced in one run. The individual reads can be assigned to the original sample afterwards via demultiplexing.

It is essential that Index Adaptor Primers i) consist of the full-length sequence with a minimum amount of truncated sequences and ii) contain an absolute minimum of other adaptor primers that will be used in the same sequencing run. Cross-contamination can cause mis-assignments of sequencing reads to other samples and can subsequently lead to wrong data interpretations.[1]

Oligonucleotide suppliers typically synthesize NGS oligonucleotides on commercial synthesizers using the phosphoramidite approach on a solid support. Afterwards, the oligonucleotides are purified by High Performance Liquid Chromatography (HPLC) to remove truncated sequences. Because traditional HPLC purification generates unacceptable levels of cross-contamination,[2] oligonucleotide vendors typically use a dedicated synthesis and HPLC purification pipeline which increases both manufacturing costs and delivery times. Furthermore, traces of cross-contamination cannot be fully excluded with this approach, as the long NGS oligonucleotides are very difficult to elute from HPLC columns and devices. Usual acceptance levels on the market for cross-contamination for Illumina sequencing are in the range of < 0.1% to 0.25%.[3,4]

In this application note, a new oligonucleotide synthesis and purification method is applied to produce high-quality NGS oligonucleotides, minimizing potential cross-contamination.

Table 1: Most commonly used NGS oligonucleotide types with assessment of the quality requirement. *The length range varies depending on whether adaptors are introduced into a sequencing library using a single-step approach as full primers or whether adaptors are introduced by initial ligation of short universal adaptors followed by a PCR with indexed linker primers.

Class	Type	Typical Length [nt]	Quality Requirement	Applications
Index Adaptor Primers (Full primer or linker primer)	Unique Dual Index (UDI)	50-100*	+++	All
	Combinatorial Dual Index (CDI)	50-100*	+++	All
	UDI-UMI	50-100*	+++	All
Amplification Primers		15-30	++	Microbiome, Genotyping
Sequencing Primers		15-25	++	All
Target Capture Probes		variable	+	Capture
Blocking Oligos		15-20	+	Capture

Results

In total, 14 NGS Adaptor primer pairs with a sequence length of 68 or 72 bases were synthesized at Eurofins Genomics EU using its proprietary synthesis technology and purification method. The same primer pairs for NGS were ordered from a competitor.

Each primer contained a unique index sequence, the adapter sequence, as well as a specific primer region for the amplification of a target sequence (Lambda phage). These primers were used to prepare amplicon libraries with a size of ~ 400 bp. The resulting libraries were confirmed to be the correct size following analysis using an Agilent Fragment Analyzer (Figure 1A). The analysis showed a vast number of full-length sequences demonstrating no significant amount of truncated library products. The libraries were sequenced on an Illumina NovaSeq 6000 with 150 paired-end read mode with each 400,000 planned read pairs. On average 435,000 read pairs were obtained (**Figure 1B**). The sequencing yield was uniform and within the typically observed variation. The number of reads with unexpected dual index combinations, not matching the combinations of the loaded libraries, was retrieved from the respective “undetermined.fastq.gz” file. For each possible dual index combination, the number of reads was determined (**Table 2**). Within the complete UDI set, 99.9986% of reads (range 99.9973 to 99.9995% for the individual UDI combinations) were correctly assigned to the expected unique dual index combination. Therefore, the total potential cross-contamination within the set of 14 UDI primer pairs was 0.0014%. The 14 primer pairs synthesized with a competitor were tested under identical conditions. In contrast to Eurofins Genomics EU primer pairs, they exhibited a tenfold higher cross-contamination rate of 0.0208% (Figure 2), showing maximum values up to 0.04% for specific UDI combinations.

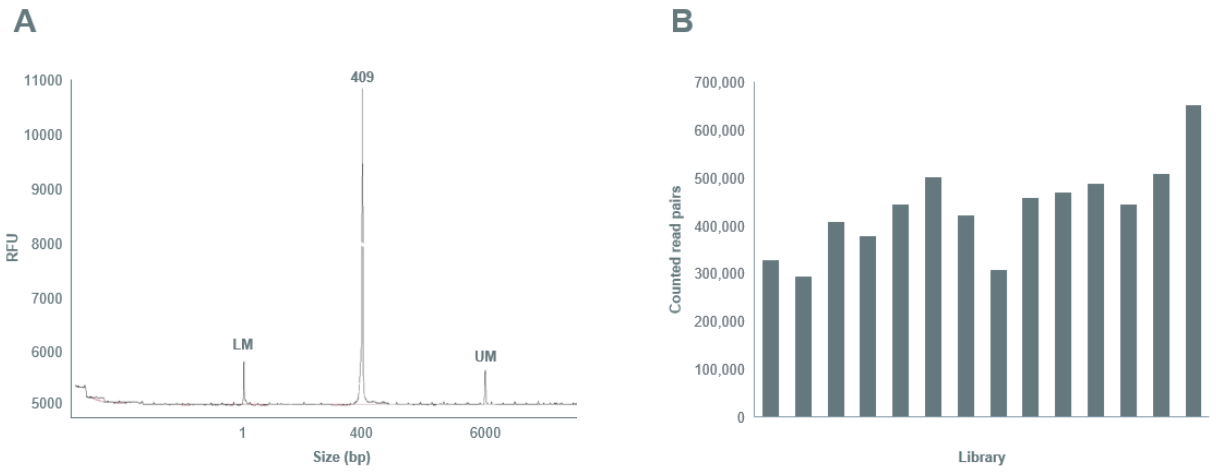
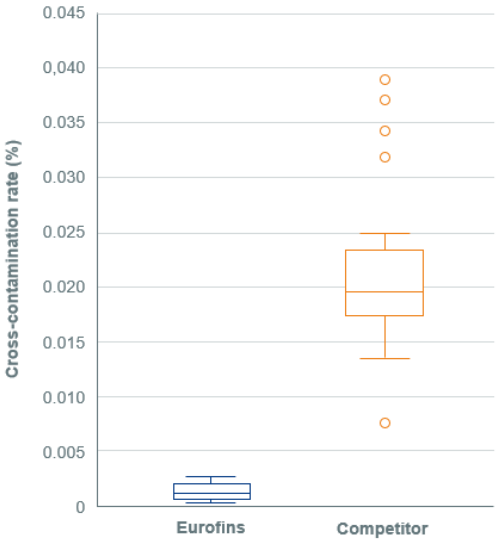


Figure 1. Library Preparation and Sequencing Results. A: Fragment Analyzer profile of the generated amplicon library; B: Sequencing yield of the 14 generated amplicon libraries.

Table 2: Example of cross-contamination measurement using 14 libraries with unique i7 and i5 indices. The number of reads for each dual index combination is given. On average 435,000 read pairs per UDI combination generated. The UDI combinations of the loaded libraries are shaded in grey.

	Index 1	Index 2	Index 3	Index 4	Index 5	Index 6	Index 7	Index 8	Index 9	Index 10	Index 11	Index 12	Index 13	Index 14
Index 1	328,370	0	1	0	0	1	0	0	1	1	0	0	0	0
Index 2	0	292,067	0	0	0	0	1	0	0	1	1	1	0	0
Index 3	1	1	407,062	1	0	0	0	1	1	1	1	0	0	3
Index 4	0	0	0	377,857	0	0	0	1	0	0	1	0	0	0
Index 5	0	0	0	1	444,754	2	0	2	1	1	0	1	1	0
Index 6	0	0	0	0	0	501,173	1	0	1	1	0	1	1	1
Index 7	0	0	0	0	1	0	419,906	0	1	0	0	1	0	1
Index 8	1	2	0	0	0	0	1	306,659	0	0	2	0	0	0
Index 9	0	0	0	1	1	0	0	0	457,550	2	0	0	0	0
Index 10	0	1	0	3	0	0	0	0	0	468,835	0	0	0	1
Index 11	1	1	1	0	0	1	1	0	3	1	486,619	0	1	0
Index 12	2	0	0	1	0	0	2	0	1	1	1	444,690	0	1
Index 13	0	1	1	3	0	2	0	0	0	1	1	0	507,643	1
Index 14	0	0	0	0	0	0	1	0	0	2	0	0	0	651,219

Figure 2. Detected cross-contamination rate using the Eurofins Genomics EU primers with the new synthesis technology compared to primers for NGS of a competitor.



Discussion

Conventional oligonucleotide synthesis methods produce a significant fraction of truncated oligonucleotides. Such by-products can lead to reduced performance in NGS applications. Consequently, oligonucleotides for NGS applications, especially for library preparation and indexing are commonly synthesized with subsequent HPLC purification. However, HPLC purification is time consuming and can lead to cross-contaminated oligonucleotides.

To produce these high performance NGS oligos, Eurofins Genomics EU utilizes a new and proprietary synthesis device, which enables best-in-class coupling efficiencies leading to near non-truncated, full-length oligonucleotides. To date, thousands of these oligonucleotides have been successfully tested in our NGS service lab for different applications.

Combined with an optimized downstream process, the risk of contamination has been reduced to a minimum.

Association of reads with other unexpected UDI primer pairs is a result of primer cross-contamination and index hopping during sequencing. Therefore, the actual cross-contamination of primers is even lower than the reported cross-contamination of 0.0014%.

Conclusion

In conclusion, the synthesized high-quality oligonucleotides can be used for any NGS application to deliver optimal, contamination free data. In addition, the high synthesis speed will revolutionize the applicability and production capacity of NGS oligonucleotides.

Learn more

The new NGS oligos are available as NGSgrade 2.0 for research purposes and as industrial-grade NGS oligos for industrial needs.

To learn more about the use for research purposes, visit https://eurofinsgenomics.eu/ngsgrade-oligos/

To learn more about the use for industrial purposes, visit https://eurofinsgenomics.eu/industrial-grade-ngs-oligos/

References

1: https://emea.illumina.com/content/dam/illumina-marketing/documents/products/whitepapers/index-hopping-white-paper-770-2017-004.pdf?linkId=36607862

2: https://www.sigmaaldrich.com/DE/de/technical-documents/technical-article/genomics/next-gen-sequencing/traditional-hplc-is-incapable

3: https://www.eurogentec.com/en/catalog/ngs-dna-base-cartridge-purification-10-nmol~ac12fd0e-07ca-4df9-aec9-73365d6a9b2f

4: https://www.sigmaaldrich.com/DE/de/technical-documents/technical-article/genomics/next-gen-sequencing/custom-next-gen-sequencing-oligos