Measuring error rates of high-fidelity DNA polymerases by NGS

Determining fidelity of Phusion Plus and Phusion Hot-Start II DNA Polymerases

Introduction

High-fidelity DNA polymerases are widely used in applications such as cloning and site-directed mutagenesis to help ensure replication of template sequences without introducing errors. Preserving accuracy of DNA sequences when amplifying the template by PCR is important for such applications because one misincorporated nucleotide may change a codon, resulting in an incorrect amino acid and possible dysfunction of the encoded protein due to misfolding. Researchers can help minimize troubleshooting in downstream applications by using a higher-fidelity DNA polymerase, and can benefit from comparing reported values of polymerase fidelity when selecting enzymes.

Several methods are available to measure polymerase fidelity, such as blue-white colony screening assays; denaturing gradient gel electrophoresis assays; and beads, emulsion, amplification, and magnetics (BEAMing) assays [1-4]. In recent years next-generation sequencing (NGS) has become a powerful method in this regard, since it can theoretically generate sufficient data to detect very low numbers of errors introduced by ultrahigh-fidelity DNA polymerases. However, in practice, the ability to detect rare errors by NGS is often obscured by a high level of background errors introduced during library preparation and/or sequencing reactions.

NGS approaches for determining polymerase fidelity should include single-molecule consensus sequencing or molecular-barcoding techniques, to identify and ignore background errors [5,6]. Once determined, the error rate of a high-fidelity DNA polymerase is usually normalized to the error rate of *Taq* DNA polymerase measured using the same assay, to account for differences in assay methodology, template DNA, and reporting units.

Here, we report a molecular-barcoding method for determining polymerase fidelity. This approach introduces unique molecular identifiers (UMIs) during MuA fragmentation in NGS library preparation. Using this method, we determined the error rates of Thermo Scientific™ Phusion™ Plus DNA Polymerase, Thermo Scientific™ Phusion™ Hot-Start II DNA Polymerase, and *Taq* DNA polymerase.

Phusion high-fidelity DNA polymerases

The family of Phusion high-fidelity DNA polymerases displays low error rates during PCR and includes over a dozen various formats. The latest addition, Phusion Plus DNA Polymerase, distinguishes itself from the rest of the family with a universal primer annealing feature, which allows primers to anneal at 60°C due to its specially formulated buffer. This feature simplifies the optimization step of primer annealing, as it does not require the annealing temperature to be determined separately for each primer set.

ThermoFisher SCIENTIFIC

Materials and methods

- Phusion Plus DNA Polymerase (**Cat. No. F630S**)
- Phusion Hot-Start II DNA Polymerase (**Cat. No. F549S**)
- *Taq* DNA polymerase (5 U/µL)
- Thermo Scientific™ 10 mM dNTP Mix (**Cat. No. R0191**)
- Thermo Scientific™ TopVision™ Agarose Tablets (**Cat. No. R2801**)
- Thermo Scientific™ GeneRuler™ 1 kb DNA Ladder (**Cat. No. SM0311**)
- Plasmid DNA template
- Forward and reverse primers
- Thermo Scientific™ MuA Transposase (**Cat. No. F750**)
- Invitrogen™ Collibri™ ES DNA Library Prep Kit for Illumina™ Systems, with CD indexes (**Cat. No. A38605024**)
- Invitrogen™ Collibri™ Library Amplification Master Mix (**Cat. No. A38539050**)
- Invitrogen™ Collibri™ Library Quantification Kit (**Cat. No. A38524100**)
- Agilent™ 2100 Bioanalyzer™ Instrument and High Sensitivity DNA Kit

Generation of PCR amplicons

A 3.75 kb sequence was amplified from a plasmid DNA template by PCR using Phusion Plus DNA Polymerase, Phusion Hot-Start II DNA Polymerase, and *Taq* DNA polymerase, following the recommended protocols. The reaction conditions and cycling protocols are provided in Tables 1 and 2, respectively.

Table 1. PCR reaction conditions and reagent amounts (for 50 µL reactions).

Component	Phusion Plus DNA Polymerase	Phusion Hot-Start II DNA Polymerase	*Taq* DNA polymerase
Reaction buffer	1X	1X	1X
Forward and reverse primers	0.5 µM each	0.5 µM each	0.5 µM each
dNTP mix	200 µM each	200 µM each	200 µM each
Template (plasmid DNA)	1 ng	1 ng	1 ng
Enzyme	0.5 µL	0.5 µL	0.25 µL
Water	To 50 µL	To 50 µL	To 50 µL

Table 2. PCR cycling protocols for the 3.75 kb target sequence.

	Phusion Plus DNA Polymerase			Phusion Hot-Start II DNA Polymerase			*Taq* DNA polymerase		
Cycling step	Temperature	Time	Cycles	Temperature	Time	Cycles	Temperature	Time	Cycles
Initial denaturation	98°C	30 sec	1	98°C	30 sec	1	94°C	2 min	1
Denaturation	98°C	5 sec	16	98°C	5 sec	16	95°C	15 sec	16
Annealing	60°C*	10 sec		65°C**	10 sec		59°C	30 sec	
Extension	72°C	80 sec		72°C	80 sec		68°C	3 min	
Hold	4°C	Indefinitely		4°C	Indefinitely		4°C	Indefinitely	

* No calculation required due to the universal annealing feature.

** Calculated with the T_m calculator at **thermofisher.com/tmcalculator**.

Preparation of sequencing libraries

The PCR amplicons obtained were fragmented with a MuA transposase. UMIs of 12 random nucleotides were introduced during fragmentation to individually tag each amplicon (Figure 1). The tagged DNA fragments were purified using the cleanup kit from the Collibri ES DNA Library Prep Kit, then amplified with the Collibri Library Amplification Master Mix. The quantity and quality of the libraries generated were assessed on an Agilent 2100 Bioanalyzer Instrument using the High Sensitivity DNA Kit, as well as by real-time PCR using the Collibri Library Quantification Kit.

Determination of polymerase fidelity

The prepared libraries were sequenced on an Illumina™ MiSeq™ instrument. Their reads were aligned to the reference sequence and grouped according to their UMI sequences. Grouping by UMI families allows differentiation of errors introduced by the polymerase during template generation from background errors introduced during library preparation and sequencing (e.g., PCR indexing, cluster generation, and sequencing reads). Errors were accounted for only when they were present in all reads in a UMI family as they must have originated during the same template generation step; otherwise, the errors were discarded as background errors. To evaluate polymerase fidelity, error rates of Phusion Plus and Phusion Hot-Start II DNA Polymerases were normalized to the error rate of *Taq* DNA polymerase.

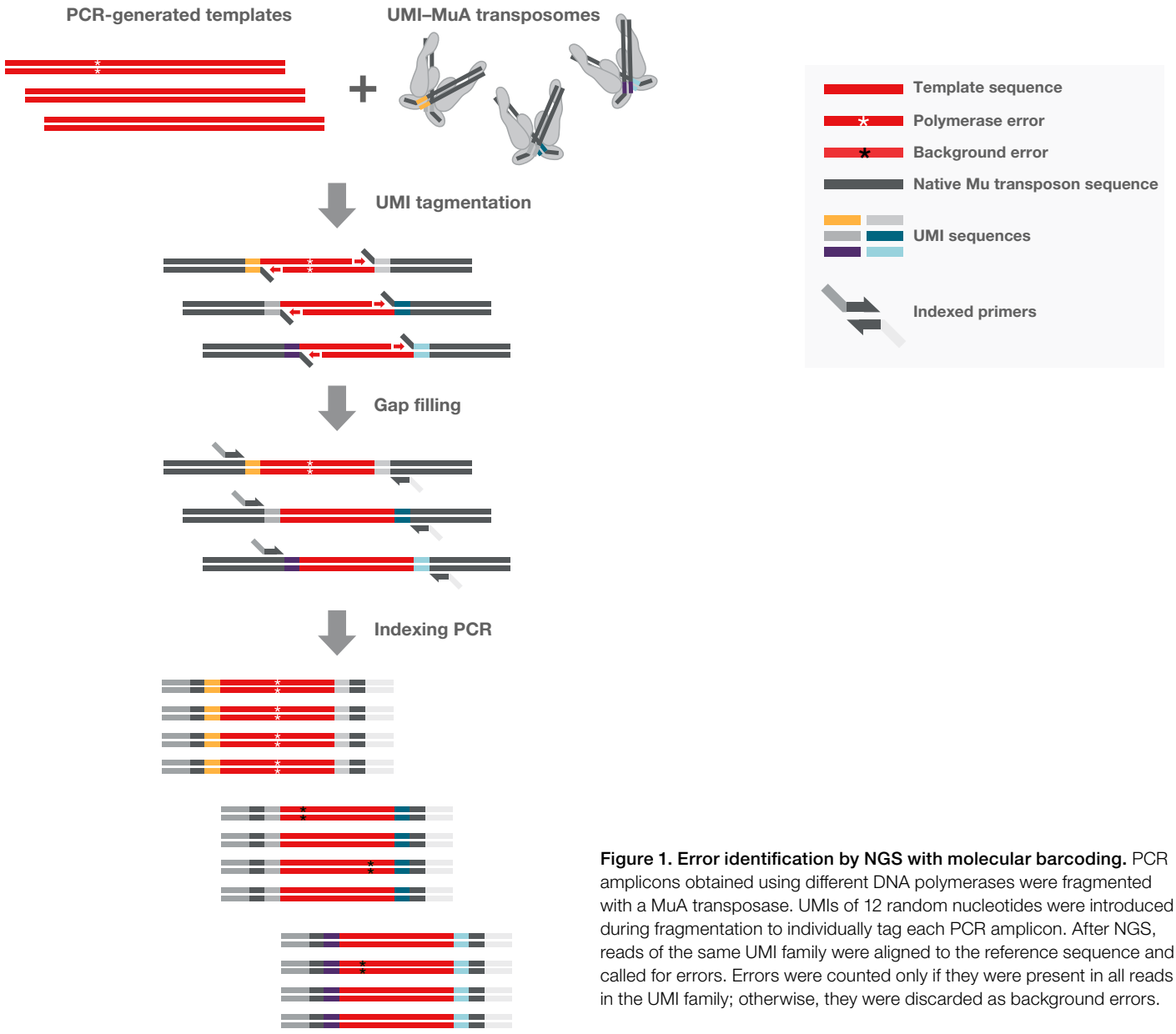

Figure 1. Error identification by NGS with molecular barcoding. PCR amplicons obtained using different DNA polymerases were fragmented with a MuA transposase. UMIs of 12 random nucleotides were introduced during fragmentation to individually tag each PCR amplicon. After NGS, reads of the same UMI family were aligned to the reference sequence and called for errors. Errors were counted only if they were present in all reads in the UMI family; otherwise, they were discarded as background errors.

thermo scientific

Results

Measured by molecular-barcoding NGS, the error rates of Phusion Hot-Start II DNA Polymerase and *Taq* DNA polymerase are 1.58 x 10^{-6} and 7.77 x 10^{-5}, respectively. These values vary from their previously reported error rates due to differences in assays and methodologies used. To account for these variations, the error rate of Phusion Hot-Start II DNA Polymerase was normalized to the error rate of *Taq* DNA polymerase from our NGS method, resulting in 49x fidelity of *Taq* DNA polymerase (Figure 2). This relative fidelity value is in the range of those reported using other methods: 38x *Taq* fidelity determined by single-molecule sequencing [3] and 54x *Taq* fidelity determined by BEAMing assay [4].

The error rate of Phusion Plus DNA Polymerase is on the order of 10^{-7}, which is much lower than the error rate of Phusion Hot-Start II DNA Polymerase, corresponding to >100x fidelity of *Taq* DNA polymerase (Figure 2). It is important to note that due to the extremely low number of errors at background level, it is difficult to accurately determine exact values in a statistically significant manner for DNA polymerases with >100x the fidelity of *Taq* DNA polymerase. It is possible that the unique reaction buffer of Phusion Plus DNA Polymerase, which is formulated for universal primer annealing, also improves incorporation of correct nucleotides by the enzyme, resulting in the higher fidelity of Phusion Plus DNA Polymerase.

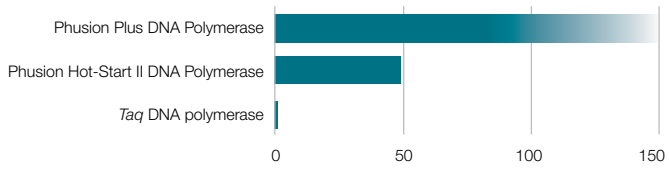

Figure 2. Fidelity of Phusion Plus and Phusion Hot-Start II DNA Polymerases relative to *Taq* DNA polymerase. Error rates were determined by NGS using molecular barcoding, then normalized to that of *Taq* DNA polymerase.

Conclusions

This technical note shows how an NGS approach was used to measure error rates of Phusion Plus and Phusion Hot-Start II DNA Polymerases. We found that the fidelity of Phusion Plus DNA Polymerase is >100x that of *Taq* DNA polymerase. We also confirmed the fidelity of its predecessor Phusion Hot-Start II DNA Polymerase to be ~50x that of *Taq* polymerase, similar to values previously measured by other methods.

References

1. Cline J, Braman JC, Hogrefe HH (1996) PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucleic Acids Res* 24(18):3546–3551.
2. Hestand MS, Van Houdt J, Cristofoli F et al. (20016) Polymerase specific error rates and profiles identified by single molecule sequencing. *Mutat Res* 784–785:39–45.
3. Potapov V, Ong JL (2017) Examining sources of error in PCR by single-molecule sequencing. *PLoS One* 12(1):e0169774.
4. Li M, Diehl F, Dressman D et al. (2006) BEAMing up for detection and quantification of rare sequence variants. *Nat Methods* 3(2):95–97.
5. Sloan DB, Broz AK, Sharbrough J et al. (2018) Detecting rare mutations and dna damage with sequencing-based methods. *Trends Biotechnol* 36(7):729–740.
6. Salk JJ, Schmitt MW, Loeb LA (2018) Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. *Nat Rev Genet* 19(5):269–285.

Find out more at **thermofisher.com/phusionplus**

ThermoFisher SCIENTIFIC