

METHODS MANUSCRIPT

qPCR-based characterization of DNA fragmentation efficiency of Tn5 transposomes

Vera Rykalina^{1,2,3}, Alexey Shadrin¹, Hans Lehrach^{1,3} and Tatiana Borodina^{1,*}

¹Alacris Theranostics GmbH, Berlin 14195, Germany; ²Department of Biology, Chemistry and Pharmacy, Free University of Berlin, Berlin 14195, Germany; and ³Vertebrate Genomics Emeritus Group, Max-Planck Institute for Molecular Genetics, Berlin 14195, Germany

*Correspondence address. Alacris Theranostics GmbH, Berlin 14195, Germany. Tel: +49-30-8431-225-25; Fax: +49-30-8431-225-41; E-mail: t.borodina@alacris.de

Abstract

Here, we describe an electrophoresis free assay for characterizing Tn5 transposomes fragmentation efficiency in a tagmentation reaction, in which double-stranded DNA is fragmented and tagged with adapter sequences. The assay uses plasmid DNA as a reference tagmentation substrate. Fragmentation efficiency is analyzed by comparative qPCR which measures the difference (Δ Ct) in amplification of a specific plasmid region before and after tagmentation: more efficient fragmentation is characterized by a larger number of cleavage events within the amplified region, a delayed increase in the amplification curve and as a result, a larger Δ Ct. Tagmentation reactions characterized with the same Δ Ct exhibit the same fragment size distribution on an agarose gel. The Δ Ct values measured can be used to quantitatively determine the relative performance of Tn5 transposome assemblies in optimization experiments and to standardize between batch variations in transposomes for use in next-generation sequencing library preparation. Moreover, the use of a reference tagmentation template added during next-generation sequencing library preparation enabled monitoring of the input DNA fragmentation. The presented qPCR-based assay is quick, contamination-safe, high-throughput and cost-efficient.

Keywords: Tn5 transposase; fragmentation efficiency; transposomes; tagmentation; efficiency; comparative qPCR

Introduction

Over the last decade, there has been increasing demand for the hyperactive variant of the Tn5 enzyme due to its use in next-generation sequencing (NGS) applications. Its ability to fragment DNA and at the same time attach adapter sequences to the ends of the fragments (tagmentation reaction) considerably shortens NGS library preparation procedures and decreases the amount of DNA required due to the omission of several purification steps. The original Tn5-based approach for genomic DNA library sequencing [1] has been quickly adapted to other NGS applications, for example, RNA-Seq [2, 3], bisulfite sequencing [4, 5] and chromatin analysis [6–8]. Despite the recognized benefits of the Tn5 enzyme, high cost and restricted availability

remain key barriers to the wider use of transposase-based NGS protocols and further technological developments in this area. For example, the Tn5 transposase used in the Illumina NGS library preparation kits (e.g. Nextera DNA Library Preparation Kit, #FC-121-1031) is preassembled with oligonucleotide adapters and the complexes are very stable making it impossible to flank library fragments with user-determined sequences. The pure Tn5 enzyme from Epicentre EZ-Tn5 insertion kits can be used for in-house transposome assembly [5], but the costs are prohibitive for larger-scale use. Most recently, however, Picelli et al. [9] have published detailed procedures for producing Tn5 transposase, which compares favorably with commercially available products for NGS library preparation.

Received: 17 October 2016; Revised: 6 January 2017; Editorial decision: 9 January 2017; Accepted: 26 January 2017

© The Author 2017. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

In house production of the Tn5 transposase will increase its availability, facilitating its adaptation for various applications and for studying its properties [9, 10]. When we started to work with in house produced Tn5 transposase for NGS applications, it became apparent that there is no easy way to determine enzyme activity in order to ensure performance compatibility between batches. Classical determination of activity relates to the ability of transposase to excise a transposon from a donor molecule, for example, the unit of activity definition: “One unit of EZ-Tn5 Transposase catalyzes the release of the donor backbone fragment from 1 mg of transposed DNA in 1 hour at 37°C, as determined by agarose gel electrophoresis” (EpiCentre EZ-Tn5 insertion kit). For NGS applications, however, this definition is not suitable because the transposase does not need to excise the transposon from a template. The working units are not transposase molecules but transposomes—stable complexes of two transposase molecules and two oligonucleotide adapters containing double-stranded mosaic ends (MEs) sequences for transposase binding—which cut DNA and ligate adapter sequences to the 5'-ends of the fragments.

Currently, fragmentation efficiency of tagmentation is assessed by visual analysis of the size distribution of the resulting DNA fragments on an agarose gel or BioAnalyser. The need for a quick, high-throughput method for estimation of tagmentation efficiency has been highlighted previously [11]. Bogdanoff and co-workers tried to establish a comparative qPCR method amplifying the whole range of fragments; however, they demonstrated that qPCR was unable to resolve libraries of differing fragment size distribution.

Here, we present a robust, electrophoresis-free method for evaluating the fragmentation activity of Tn5 transposomes. Our approach involves a reference tagmentation template which is used alone or with a carrier DNA. Fragmentation efficiency is estimated relative to a non-fragmented DNA control using comparative qPCR with primers amplifying specific regions on the reference template. ΔC_t during amplification of a PCR product before and after tagmentation is highly reproducible for transposome assemblies and serves as a comparison value in our assay. We use this assay for normalizing variations in activity exhibited by different transposomes batches, for determining the influence of reaction conditions on tagmentation and for controlling tagmentation of genomic DNA samples.

Materials and methods

Tn5 preparation

pTXB1-Tn5 was a gift from Rickard Sandberg (Addgene plasmid #60240). The Tn5 in this construct is fused with intein-chitin binding domain (CBD), which enables efficient purification of the protein using a specific affinity matrix. The domain is then cleaved releasing pure Tn5 enzyme.

The stab culture was streaked onto a LB agar, 100 µg/ml Ampicillin plate and grown overnight at 37°C. Several colonies were picked and grown in 2 x 5 ml of LB medium with 100 µg/ml Ampicillin overnight at 37°C, for creating pTXB1-Tn5 reserve glycerol stocks and for plasmid purification. Further transformation into an *Escherichia coli* expression strain and protein production were performed according to Picelli et al. (2014), with minor modifications. Insertion of Tn5 was verified by Sanger sequencing and clones were transformed into T7 Express *lysY/lq* Competent *E. coli* (NEB, #C30131), according to the manufacturer's instructions. One colony per clone was grown in 50 ml LB with 100 µg/ml Ampicillin overnight at 37°C. Ten milliliters of

the overnight culture were used to inoculate 1 l of LB medium with 100 µg/ml Ampicillin. The culture was grown to OD 0.5 and induced with 250 µl of 1 M IPTG for 4 h at 23°C. Bacteria were harvested by centrifugation (5000 g, 4°C, 15 min) and the pellet washed with 80 ml of cold HEX buffer (20 mM HEPES-KOH, pH 7.2, 0.8 M NaCl, 1 mM EDTA, 0.2% Triton X-100) and frozen at -80°C overnight. Pellets were thawed on ice, resuspended in 40 ml of cold HEGX buffer (HEX buffer, 10% Glycerol) with 1% cOmplete protease Inhibitor Cocktail (Sigma, #4693132001) and lysed using a French Press (~80 psi, 4 x 10 presses, circulating). Cell debris was removed by centrifugation at 14,000 x g for 30 min at 4°C. *Escherichia coli* genomic DNA was precipitated by adding 2.4 ml of 10% polyethyleneimine (Sigma, #P3143) to the supernatant followed by centrifugation at 15,000 x g for 10 min at 4°C. Eighty milliliters of supernatant was mixed with Chitin Magnetic beads (NEB, #E8036L) resuspended in 10 ml cold HEGX buffer; the beads were previously washed in 100 ml of HEGX buffer. Binding was carried out for 4.5 h at 4°C on a rotator. Beads were collected on a magnetic stand and washed 5 x (5 min incubation between each wash) with 56 ml of HEGX buffer at 4°C. Finally, the beads were resuspended in 20 ml of cold HEGX buffer with 50 mM DTT and incubated for 36 h at 4°C on a rotator. During this step, the intein-CBD domain is cleaved from the fusion protein and pure Tn5 is released into the solution. The beads were again collected on a magnetic stand and the supernatant transferred to a new tube. Beads were washed 5 x with 8 ml of cold HEGX buffer; the eluate was transferred to a separate tube between each wash. Tn5 concentration was measured in each supernatant using the Qubit® Protein Assay kit (Thermo Fischer Scientific, #Q33211) and protein-containing samples were combined. Protein concentration and buffer exchange to 2x Tn5 Exchange buffer (100 mM HEPES-KOH, pH 7.2, 0.2 M NaCl, 0.2 mM EDTA, 0.2% Triton X-100, 2 mM DTT) with 20% glycerol were performed using Amicon® Ultra-4 Centrifugal Filter Devices (Millipore). The Tn5 solution was then mixed at a ratio of 1:1 with 80% glycerol to provide 50% glycerol in buffer for storage. From 1 l of induced bacterial culture, we obtained approximately 1 ml of stock Tn5 solution at a concentration of 40 µg/µl or 750 pmol/µl (Tn5 MW = 53300).

Transposomes

For transposome assembly, a 20 µM Tn5 solution was prepared in 1x Tn5 Exchange buffer with 50% glycerol. The same aliquot of Tn5 solution was used for all transposome assemblies described in this paper.

The Tn5 transposon end adapters used in this work are the Illumina NGS libraries preparation scheme adapters. Oligonucleotides were purchased by TIB Molbiol (Berlin, Germany): #Tn5ME-A 5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3', #Tn5ME-B 5'-GTCTCGTGGCTCGGAGATGTGTAT AAGAGACAG-3', #Tn5MErev 5'-CTGTCTCTTATACACATCT-3'. Underlined regions correspond to the double-stranded part of the adapter—ME sequence, recognized by the transposase.

80 µM #Tn5ME-A and 80 µM #Tn5ME-B were annealed to 80 µM #Tn5MErev at a ratio of 1:1, to generate 40 µM Tn5ME-A and Tn5ME-B adapter stocks. Stocks of 20 µM Tn5ME-A/B adapters (each adapter 10 µM), 50% glycerol solution were prepared. For transposome preparation, equal volumes of 20 µM Tn5 enzyme and 20 µM Tn5ME-A/B adapters were mixed and incubated for 1 h at room temperature. The resulting assembly was stored at -20°C until use. For evaluation of the in house produced Tn5 transposomes, the TDE1 Tagment DNA enzyme from Illumina Nextera Rapid Exome kit (Illumina, #FC-140-1006) was used.

In house transposome assemblies were added to tagmentation reactions without removal of residual adapters and free Tn5 molecules. It was therefore not possible to determine the actual concentration of transposomes. In the Illumina TDE1 mixture, the concentration of transposomes is also not indicated. Therefore, for both the in house prepared and commercial mixtures, the volume was used as a measure of the amount of the transposome assembly used in tagmentation reactions.

Preparation of standard tagmentation template

pUC19 plasmid was purchased from NEB (#N3041S). The 3 kb plasmid was linearized with EcoRI restriction endonuclease (Invitrogen, # 15202-013): 1 µg of plasmid was incubated with 15 u of EcoRI in 20 µl of 1×Buffer H at 37°C for 2h. The enzyme was heat inactivated at 65°C for 20 min and the reaction purified with the QIAquick PCR Purification Kit (Qiagen, # 28104). Digestion efficiency was checked on a 1% agarose gel and DNA concentration measured on a Qubit Fluorometer using the Qubit® dsDNA BR Assay Kit (Thermo Fischer Scientific, # Q32850); 22 ng/µl pUC19/EcoRI stock was used in all described reactions.

Tagmentation

Tagmentation reactions were performed in 1x TB Buffer (10 mM TrisCl pH 8.0, 5 mM MgCl₂, 10% DMF (Sigma, # D4551) with 50 ng of pUC19/EcoRI in 25 µl. For the spike-in experiment, 0.5 ng of the reference plasmid was added to 50 ng of human genomic DNA (Bioline, # 35025) for tagmentation.

For evaluation of the fragmentation efficiency assay reproducibility and tagmentation bias (Figs 2 and 3), first the transposome assembly dilutions were prepared from the stock assembly in 0.5x Storage Buffer and 50% Glycerol and then equal volumes of transposomes were added to the tagmentation reactions run in parallel. This ensured equal buffer composition (salt, glycerol) in all reactions. A negative tagmentation control without transposase was always performed in parallel.

Reaction temperature varied depending on the experiment. For evaluation of the fragmentation efficiency assay reproducibility and tagmentation bias (Figs 2 and 3), tagmentation was performed for 1h at 37°C. This incubation period was selected to attenuate differences related to handling of tubes, which may have more influence over shorter time periods. For the comparison of the in house Tn5 transposomes and Illumina TDE1 mixture and for tagmentation of genomic DNA with a plasmid spike-in (Figs 4 and 5), tagmentation reactions were performed for 10 min at 58°C, using conditions recommended by Illumina (Nextera protocols for the NGS library preparation). Tagmentation reactions were stopped by adding 2% SDS (to the final concentration 0.08%) and incubating for 7 min at 55°C. Different reaction conditions influence tagmentation efficiency: Δ Ct plots for the 1240 bp PCR detection region on Figs 2C and 4A characterize the same transposome assembly.

Tagmentation reactions were purified with Agencourt AMPure XP Beads (Beckman Coulter, # A63880), which provide efficient removal of DNA molecules of smaller sizes (cut-off is adjustable). Purification was conducted as the 10 µM adapters in the in house transposome assembly inhibit PCR reactions, especially when more than 30 pmol are taken per 25 µl of tagmentation reaction (data not shown). Purification was carried out according to the manufacturer's protocol, with the following settings: beads were added to the Tn5 reactions at a 0.8:1 ratio; beads were washed with 70% EtOH; DNA was eluted in the 10 mM TrisCl, pH 7.5, in the same volume as taken for purification.

qPCR

Aliquots of the purified tagmentation reactions were analysed by real-time PCR at the following dilutions: 1/150 for plasmid-only tagmentation; 1/25 for the plasmid spike-in experiments. All qPCR experiments were conducted on the StepOne™ Real-Time PCR machine (Thermo Fischer Scientific), with the SYBR® Green PCR Core Reagents (Applied Biosystems, # 4304886) using the following cycling conditions: 95°C for 10 min followed by 40 cycles of 95°C for 15 s and 63°C for 15 sec and 72°C for 60 s (for PCR products <1 kb) or 70 s (for 1240 bp PCR product) or 120 s (for 2248 bp PCR product). Each reaction contained 0.5 µM each of forward and reverse primers, in a final reaction volume of 20 µl. The PCR primers used and their pUC19 co-ordinates are listed in [Supplementary Table S1](#).

Visualization of fragment sizes

Fragment sizes were checked by loading half of the tagmentation reaction volume on a 1.1% UltraPure Agarose (Invitrogen) gel. Electrophoresis was performed in 1x TAE buffer at 120 V for 2.5 h. In tandem, 1 µl of the purified tagmentation reaction was checked using the Agilent 2100 BioAnalyser High Sensitivity DNA assay.

Results and discussion

Assay principle

The principle of the approach is depicted in [Fig. 1](#). Here, the fragmentation efficiency of a tagmentation reaction is estimated relative to a non-tagmented DNA control. Tagmentation is performed on a reference tagmentation template—single type double-stranded DNA molecules, in our case a ~3 kb plasmid. One sample (left column on [Fig. 1](#)) is treated with Tn5 transposomes, which bind to the molecules, insert breaks and ligate oligonucleotide adapters to the 5'-ends at fragmentation sites. The untreated negative control (right column) consists of an equal amount of template DNA and is processed identically to the treated sample—but without transposomes—so plasmid molecules remain intact. Aliquots of both reactions are then used for real-time PCR with plasmid specific primers (step 2). In the untreated sample, all DNA molecules serve as PCR templates. In the transposome-treated sample, some DNA molecules are fragmented within the amplified region and are thus excluded from amplification. The difference in the amount of amplifiable DNA in the samples results in different Ct values. As shown in the amplification plot, a signal from the tagged sample is observed at a higher cycle number than the untreated control.

Thus, instead of visualizing the distribution of fragment sizes on a gel, the tagmentation-related change in the amount of amplifiable DNA molecules is measured directly by qPCR, using a reference tagmentation template and specific PCR primers. Because the Ct value for a PCR product of a defined size is directly dependent on the amount of initial PCR template, we can gauge the efficiency of the tagmentation reaction: the more efficient the tagmentation reaction is, the more fragmentation will occur within the PCR target region (and vice versa) and the Δ Ct between the intact and the tagged template increases. Δ Ct can therefore be used to characterize the absolute change in amplification of the template, when treated and untreated samples are compared ([Fig. 1](#)). The difference in Δ Ct values between multiple tagmentation reactions allows the assessment of their relative performance.

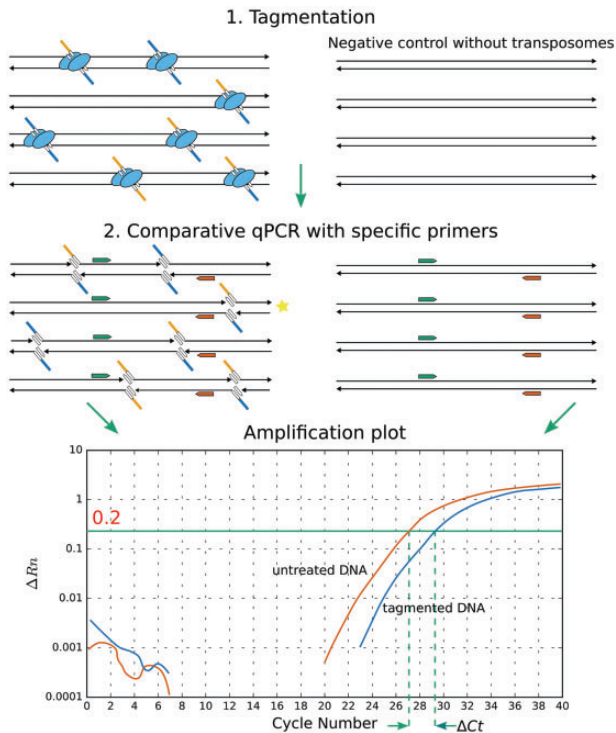


Figure 1: schematic view of the principle of the Tn5 transposomes fragmentation efficiency test. During the tagmentation (step 1) equal amounts of plasmid molecules are processed in the same conditions in parallel with (left column) and without (right column) transposomes. Transposomes are depicted as double circles, each circle with partly double arrows, corresponding to transposase dimers bound to Illumina oligonucleotide adapters. Transposase recognition sites are shown as empty double arrows and two types of single-stranded tails are colored with yellow and blue. After removal of transposomes, samples are analyzed with real-time PCR (step 2). PCR primers are plasmid-specific and shown as green and orange arrows. All molecules in the untreated sample can be amplified. In the transposome-treated sample only those molecules may be amplified which have no transposase-inserted breaks in the region between the PCR primers: from the four drawn DNA molecules only one (marked with a star) gives rise to a PCR product. The amplification curve demonstrates the difference in Ct (here two cycles) corresponding to the difference in the amount of amplifiable templates in tagged and untreated DNA samples (here four times).

Assay design

The choice of the tagmentation template and PCR detection region(s) for the fragmentation efficiency assay are influenced by a number of factors. To identify which parameters might influence the performance of the assay we estimated the expected change of the ΔCt between tagged DNA and non-tagged control depending on the fragmentation efficiency. The estimate was based on the assumptions that (1) tagmentation is truly random; (2) tagmentation can occur at any position on the plasmid, independent of other fragmentation sites and (3) during PCR amplification the amount of product doubles during each cycle. Using these assumptions the probability that k tagmentation events occur in a DNA molecule of size L outside the PCR detection region of size l is calculated as: $p = ((L-l)/L)^k$. If tagmentation is performed on N molecules, the amount of templates for PCR is $n = N^*p$. Then, the ΔCt between tagged and non-tagged DNA can be estimated as $\Delta Ct = \log_2(N/(N^*p)) = -\log_2(p) = k * \log_2(L/(L-l))$. Thus, ΔCt does not depend on the amount of templates in the tagmentation reaction and aliquot taken for PCR, they just need to be the same for the samples being compared. The sensitivity of the assay depends on

the ratio of L and l : the larger the difference between L and l , the higher the probability that additional fragmentation events will happen outside the detection region, so p will be closer to 1, and consequently, the ΔCt will be closer to 0. For example, registering fragmentation events with a 2 kb PCR region on a human genomic DNA is not suitable for revealing slight changes in fragmentation efficiency. To set up a sensitive test system, the amount of DNA outside the PCR-detected region should be minimized and, ideally, the whole tagmentation template should be amplified. The tagmentation template used for our system is the 2686 bp pUC19 plasmid DNA, which is widely used and available commercially. We wanted the template to be large enough to produce tagmentation fragments of 200–500 bp as required for typical NGS applications. For a template larger than ~3 kb long, however, it would have been difficult to select a PCR detection region close to the size of the template—in our internal tests for PCR products larger than 2400 bp changes of Ct are not proportional to the changes in the amount of starting material (data not shown). To ensure amplification of a single PCR from the plasmid template, pUC19 was linearized with EcoRI and four PCR detection regions of different lengths—310, 610, 1240, and 2248 bp (Fig. 2A) were selected. The expected ΔCt for these PCR detection regions was calculated and plotted against the amount of fragmentation events per template pUC19/EcoRI molecule (Fig. 2B). Using the assumptions mentioned above, ΔCt exhibits a linear relationship with the number of fragmentation sites and larger PCR detection regions are more sensitive to changes in fragmentation efficiency, as shown by the relative location of the four ΔCt amplification plots (Fig. 2B).

To replicate the effect of increased tagmentation efficiency, we added more transposomes to the tagmentation reaction. A series of tagmentation reactions with double dilutions of the in house Tn5 transposome assembly was performed in triplicate. Reactions were purified with Agencourt AMPure XP beads to get rid of the non-ligated transposon end adapters, and equal aliquots of each reaction were tested by qPCR with each of the four selected PCR primer pairs. To evaluate the effect of technical variation of tagmentation on Ct values, PCR variability was minimized by using a single PCR master mix per PCR detection region for all three tagmentation series replicates. Average ΔCt values between the non-tagmented and tagged plasmid for each detection region were plotted against transposome assembly amounts (Fig. 2C). As expected, depending on the size of the amplified region, the system reacted differently to the fragmentation efficiency of tagmentation. For the largest 2248 bp PCR detection region, each doubling of the amount of transposomes led to a significant increase of ΔCt ; whereas, for the smallest PCR product (310 bp), even larger dilution intervals were not reflected well in the ΔCt differential. As for the Ct deviation between amplified tagmentation replicates, it was comparable to that measured between technical PCR replicates (Supplementary Fig. S1), indicating that the tagmentation reaction is highly reproducible.

In addition to the selection of the tagmentation template and the PCR detection region lengths, we next evaluated whether the position of the PCR detection region is important for assay design and compared the performance of three different PCR detection regions of the same size on the same pUC19/EcoRI template (Fig. 3A). A centrally located region of 610 bp used for the initial experiments, plus two regions of 591 bp and 602 bp, located at either end of the linearized plasmid were selected. All three regions were subjected to the same series of tagmentation reactions; however, the change in ΔCt values was variable in all three target regions, with the centrally located

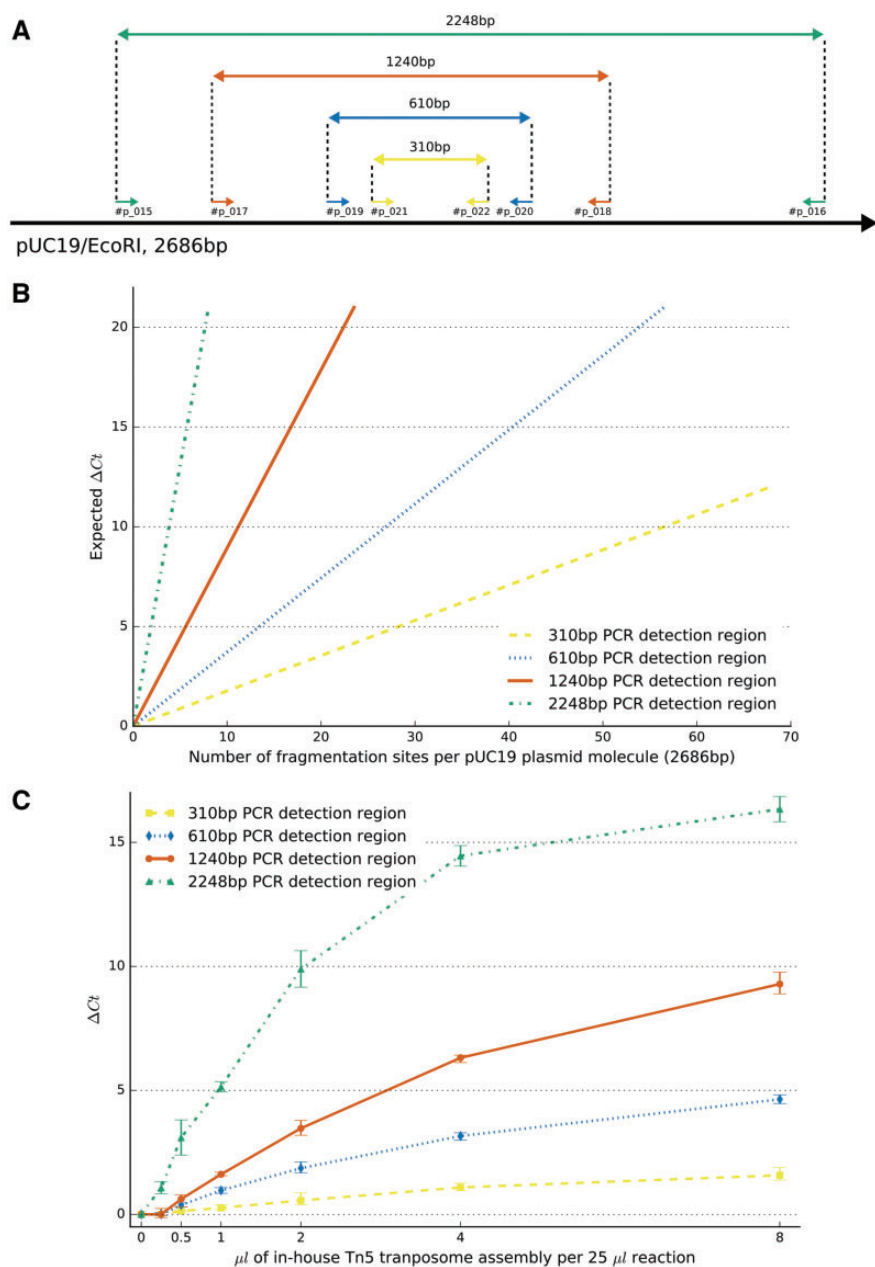


Figure 2: (A) location of the four PCR primer pairs amplifying regions of different sizes on the pUC19/EcoRI, scale preserved. The same color coding is used for plots in this figure. (B) The plot shows values of estimated ΔC_t between tagged and intact 2686 bp long pUC19/EcoRI samples for the increasing number of fragmentation sites per plasmid molecule. Four graphs correspond to four PCR detection regions of different length. The more fragmented the plasmid, the less amplifiable templates are left and the larger is ΔC_t . The larger the PCR detection region, the larger the slope. (C) Experimental ΔC_t curves obtained for the same four PCR detection regions on a series of tagmentation reactions. X-axis shows the amount of transposomes used as microliters of transposome assembly. Y-axis shows average ΔC_t obtained in qPCR for the three replicate tagmentation reactions. The correspondence of X-axes on Figs 2B and 2C would depend on the particular transposome assembly and reaction conditions used.

PCR detection region seemingly harboring more fragmentation sites than the other regions. No variations in performance was found in the non-tagmented controls, including plasmid dilutions that simulated decreasing amounts of amplification template (Supplementary Fig. S2), confirming a bias in tagmentation activity across the length of the plasmid. Tn5 is known to have a tagmentation bias close to the ends of DNA molecules [1], thus the difference in activity exhibited is most likely a feature of the PCR detection region. In accordance, analysis of BioAnalyser and agarose gel images of the tagged DNA (Fig. 4B and Supplementary Fig. S3) revealed distinct bands within

the smear of fragmented DNA, especially at lower transposomes concentrations.

The discussed parameters of qPCR-based fragmentation efficiency evaluation of tagmentation reactions should be taken into account when establishing such an assay in the laboratory.

Equalizing transposome assemblies

In our experience, the efficiency of the in house produced Tn5 transposome assemblies varies slightly even if the same protocol was used for their preparation. To provide consistency of

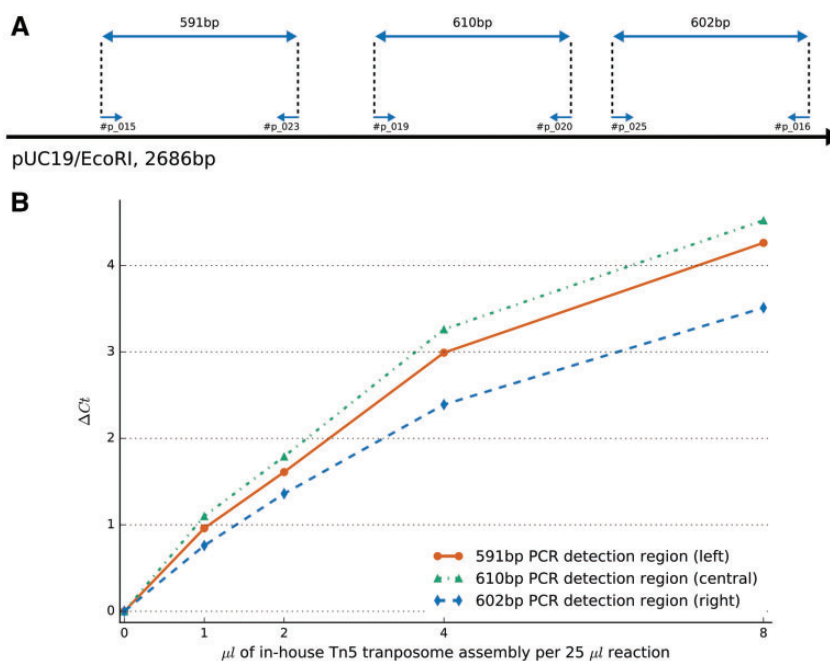


Figure 3: (A) location of the three PCR primer pairs amplifying ~600bp regions from the central and outside parts of the pUC19/EcoRI. (B) Tagmentation bias along the plasmid DNA. qPCR analysis of a series of tagmentation reactions using three non-overlapping PCR detection regions demonstrated different fragmentation efficiencies within these regions. Highest ΔC_t was obtained for the centrally located PCR detection region.

fragmentation, the activity of transposome assemblies have to be normalized relative to each other or to commercial transposomes. We also recommend controlling fragmentation activity of stored transposome assemblies, as in our experience in house Tn5 transposomes stored at -20°C gain efficiency over time, probably due to the presence of free Tn5 and oligonucleotide adapters in the solution.

We routinely use the qPCR-based fragmentation efficiency assay on plasmid DNA to characterize and compare batches of in house Tn5 transposome assemblies. Figure 4A shows ΔC_t values plotted against the amount of transposome assembly taken for tagmentation for a batch of in house transposomes (in house Tn5 transposomes assembly, red line) and the TDI transposome assembly from the Illumina Nextera exome rapid kit (Illumina TDI transposome assembly, blue line). The change of ΔC_t values with the increase of the transposomes volume in the reaction correlates well with the change of tagmentation fragments size ranges on BioAnalyser (Fig. 4B) and agarose gel images (Supplementary Fig. S3). To perform tagmentation with the same efficiency using different transposome assemblies, it is necessary, in the tagmentation reaction, to use amounts of assemblies which generate the same ΔC_t value in the qPCR assay. The ΔC_t value is a quantitative parameter; assemblies may be matched to each other using ΔC_t plots as shown in Fig. 4A (Supplementary Fig. S4). ΔC_t is also a very convenient indicator of the influence of reaction conditions on fragmentation: if the changed parameter improved the tagmentation efficiency then the ΔC_t will increase, if it inhibits the reaction the ΔC_t will decrease (Supplementary Fig. S5).

Use of the reference tagmentation template as a spike-in to the genomic DNA

Monitoring tagmentation efficiency on a standard template is convenient for establishing reaction conditions and characterizing transposome assemblies, but due to differences in DNA

complexity, the distribution range of fragment sizes obtained using plasmid DNA can't be directly extrapolated to other templates, such as high molecular weight genomic DNA. However, plasmid DNA can be added directly to the tagmentation reaction alongside the sample DNA as a "reaction reference". In the example shown (Fig. 5), visual evaluation of in house and Illumina transposomes tagmentation reactions products on an agarose gel indicated that the fragment sizes of tagged human genomic DNA corresponds to the ΔC_t obtained by qPCR monitoring of fragmentation of a 1% plasmid spike. On lanes 5 and 6, the genomic DNA smears look very similar and qPCR performed using the 1240 bp PCR detection region gave ΔC_t values of 5.08 and 4.89 cycles, respectively. On lanes 3 and 4, the smears are obviously different, correlating with the ΔC_t values of 2.05 and 1.62. For a large project, using DNA samples of the same quality, it would be most convenient to perform a gel check of the tagged input DNA only once, allowing the association between the necessary size of the input DNA and the ΔC_t of the added plasmid DNA to be established. Following this check, the DNA spike-in can be used to check the fragmentation efficiency and consistency of tagmentation performance in qPCR assay using only a tiny aliquot of the tagmentation reaction—we routinely use 1/25 of the reaction volume.

Conclusions

For reproducible work with transposase in NGS applications, it is necessary to have a robust tool for estimating fragmentation efficiency in tagmentation reactions. This is important both for routine NGS projects—to control transposome assembly efficiencies and monitor results over the project duration, and for methodological tasks to determine the effects of reaction variables.

The suggested qPCR-based fragmentation efficiency test is based on the introduction of a standard tagmentation template and fixed PCR detection region. This setup enables the use of

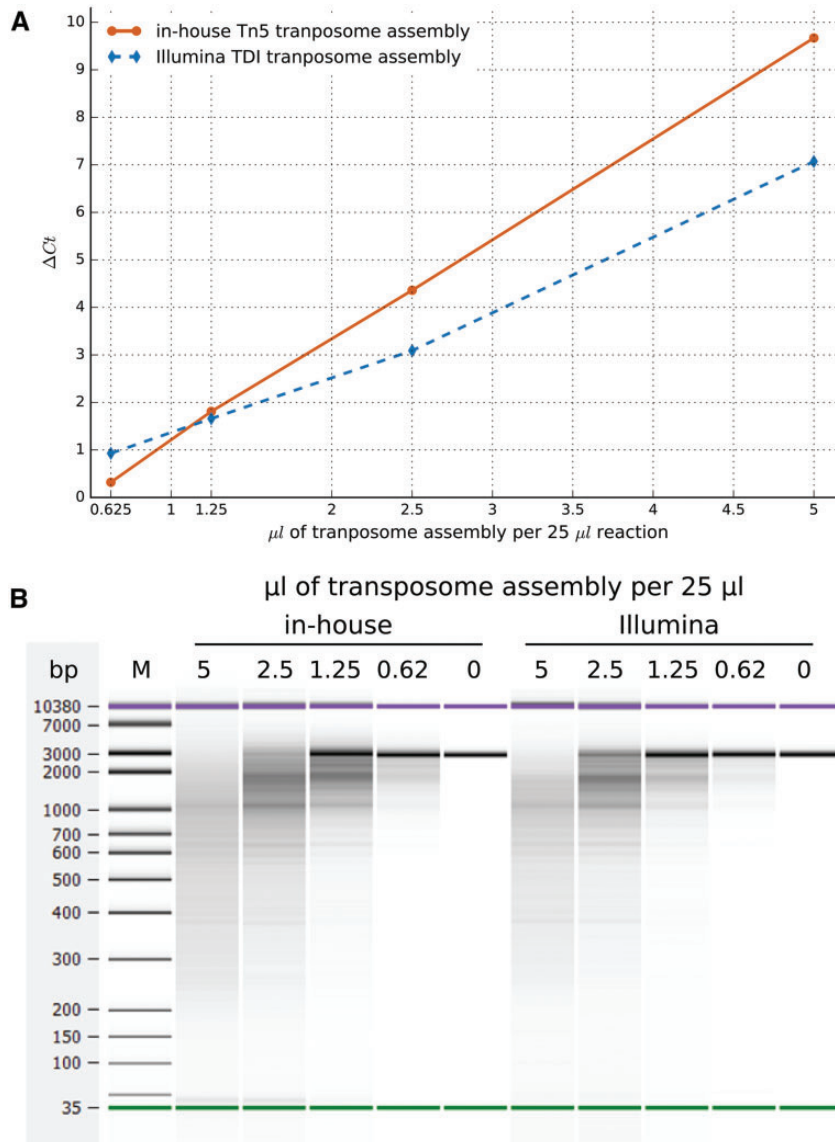


Figure 4: (A) ΔC_t graphs for fragmentation reactions performed by in-house and Illumina transposomes on 50 ng of pUC/EcoRI. qPCR was performed using 1240 bp PCR detection region (B). BioAnalyser image of fragment sizes of the same fragmentation reactions. The difference in fragments distribution between the lanes corresponds to the ΔC_t difference in the qPCR-based fragmentation efficiency assay. Prevalence of large fragments at low transposomes concentrations and noticeable brighter bands within the smear is a visual demonstration of a transposition bias.

comparative PCR, a standard molecular biology tool, and converts the task of determining the difference between fragment size distribution ranges to determining the difference between the amounts of a particular amplification template in the solution. The use of reference DNA also facilitates comparison of results obtained in different experiments. Reference template and PCR detection regions do not necessarily need to be as suggested here; although the pUC19 DNA used in this work is inexpensive and available in standard quality. For some applications, selection of another template may be more relevant; for example, when a reference template is used to spike the input DNA used for preparing NGS libraries, it would be advantageous to be able to easily discard this DNA, avoiding unnecessary sequencing. In this case, a dUTP-containing PCR product, which can be tracked for fragmentation efficiency and then destroyed by UDGase in the downstream steps, could be a reasonable alternative.

Fragmentation of the non-complex and short fragments of standard fragmentation template DNA used in this work was shown to be prone to bias; however, for characterization of the fragmentation efficiency this is not important. Since fragmentation is generally biased, it is not possible to use the absolute number of fragmentation events per PCR detection region and predict the average size of the fragmented DNA fragments. For relative comparison of fragmentation reactions even a biased but reproducible test system is adequate. What is important is the correspondence of the relative ΔC_t value and relative size distribution on the gel. Of course, the bias should be consistent; therefore, the method would not work for comparison of transposases which have different biases. In such cases, the detected region would not remain proportional to the rest of the DNA.

The method described here is qPCR-based and therefore extremely robust, quick, cost- and time-efficient. In principle, this

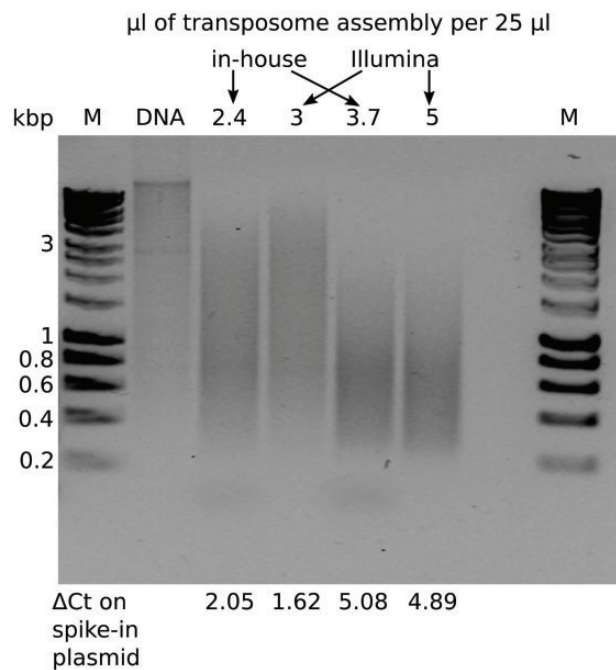


Figure 5: correlation between visual and qPCR-based fragmentation efficiency analysis. 50 ng human genomic DNA (input DNA) samples with 0.5 ng pUC/EcoRI spike-in were tagged with in house and Illumina transposomes (amounts taken for reactions are indicated above the gel image). Tagmentation reactions were purified and analyzed on an agarose gel. In parallel, 1/125 reaction aliquots were used for a qPCR efficiency test (ΔC_t values are indicated below the gel image). The amounts of in house and Illumina transposomes used for the tagmentation reactions were selected according to the comparison ΔC_t graph built on reference template on Fig. 4A, explained in detail in Supplementary Fig. S4. Correctly normalized amounts of in house and commercial transposomes give visually identical smears of genomic DNA (lanes 5 and 6). ΔC_t values for these samples are also very similar (0.19 cycles difference). Inconsistencies in normalizing results in differing smears on the gel (lanes 3 and 4) and correlated increase in difference in ΔC_t values (0.43 cycles). The low weight bands visible below the smear of genomic DNA fragmented using in house transposomes, correspond to the residual adapters.

approach of detecting fragmentation sites within a certain region may be used for assessment of other strategies where the number of these sites is characteristic of the reaction; this includes non-transposase-based fragmentation strategies, both enzymatic and physical, and also other reaction types such as ligation. For site-specific reactions, for example, restriction, our approach would even enable the use of absolute quantification of fragmentation sites to accurately measure the activity units as the percent of cut molecules in certain reaction conditions.

Acknowledgements

The authors would like to thank Dr Alisa Fuchs for help and advice in preparation of the in house Tn5 transposase. They are very grateful to Dr Lesley Ogilvie for editing suggestions and proof-reading of the manuscript.

Funding

This work was supported by the Vertebrate Genomics Emeritus Group at the Max-Planck Institute for Molecular Genetics.

Conflict of interest statement. None declared.

Supplementary data

Supplementary data is available at *BIOMAP Journal* online.

References

- Adey A, Morrison HG, Asan et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. *Genome Biol* 2010;11:R119.
- Syed F. Application of Nextera™ technology to RNA-seq library preparation. *Nat Methods* 2010;7:2010.
- Gertz J, Varley KE, Davis NS et al. Transposase mediated construction of RNA-seq libraries. *Genome Res* 2012;22:134–41.
- Adey A, Shendure J. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. *Genome Res* 2012;22:1139–43.
- Wang Q, Gu L, Adey A et al. Tagmentation-based whole-genome bisulfite sequencing. *Nat Protoc* 2013;8:2022–32.
- Buenrostro JD, Giresi PG, Zaba LC et al. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 2013;10:1213–8.
- Schmidl C, Rendeiro AF, Sheffield NC et al. ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors. *Nat Methods* 2015;12:963–5.
- Mumbach MR, Rubin AJ, Flynn RA et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat Methods* 2016;13:919–22.
- Picelli S, Björklund AK, Reinius B et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res* 2014;24:2033–40.
- Cole C, Volden R, Dharmadhikari S et al. Highly accurate sequencing of full-length immune repertoire amplicons using Tn5-enabled and molecular identifier-guided amplicon assembly. *J Immunol* 2016;196:2902–7.
- Bogdanoff D, Jou T, Lee B. Characterization of action and efficiency of Tn5 transposase through comparative qPCR. *J Exp Microbiol Immunol* 2015;19:1–6.