scientific reports

Check for updates

OPEN

Efficiency of genome editing using modified single-stranded oligodeoxyribonucleotides in human cells

Seryoung Kim[1], Yosuke Matsushita[2,3], Toyomasa Katagiri[2,3] & Hideaki Maseda[1,4]✉

Single-stranded oligodeoxyribonucleotide (ssODN) gene editing has emerged as a promising therapeutic strategy. However, further improvements in efficiency are desired for practical application. The effects of strand length and locked nucleic acid (LNA) modification on ssODN genome editing were investigated by introducing an assay cassette into the genome of HEK293T cells and measuring precise base deletions of eight bases. The introduction of LNAs into ssODNs, five pairs of LNAs at 25–35 nt from the centre and one pair at 20–25 nt, showed approximately 18-fold higher efficiency than unmodified ssODNs of the same length in the study using 70 nt ssODNs. In addition, genome editing efficiency was further improved when LNAs were introduced at the same positions as the 70 nt ssODN, which showed the highest efficiency for the 90 nt ssODN. However, in some cases, the same number of LNA modifications could conversely reduce the efficiency, and the modification positions in the ssODN method were successfully optimised in the present study. Furthermore, the oligo DNA was shown to be effective not only for deletions but also for base substitutions, with an editing efficiency of 0.63% per cell.

Genome editing technology has become indispensable in all areas of biotechnology involving living organisms, revolutionising efficient biomass production and disease treatment by modifying the genetic blueprints of organisms. In particular, the CRISPR-Cas system is currently the most actively used genome editing technology, facilitating easy editing of target genes by expressing guide RNA (gRNA), an RNA molecule that recognises the target sequence, and Cas9, a bacterially derived nuclease, in target cells[1–3]. However, several challenges must be overcome to achieve precise editing at the intended site, including the need to introduce additional nucleic acids that allow homologous recombination at the target site, the risk of off-target editing affecting sequences other than the intended site, and the potential immune response triggered by Cas proteins derived from foreign organisms[4–7]. These obstacles hinder the active use of nucleic acids in medicine, where precision is paramount. In contrast, the single-stranded oligodeoxyribonucleotide (ssODN) method used in eukaryotic genome editing is highly safe because it allows precise target sequence modification without the need for Cas proteins, thus overcoming the drawback of CRISPR in the medical field[8,9]. However, the problem of low efficiency still remains. To address this challenge, improvements have been made to increase efficiency by introducing foreign sequences at the ends of transgenic ssODNs or by modifying specific sequences[10–15]. However, most studies have focused on antisense or short-stranded nucleic acids, and very few examples of sense ssODNs of approximately 100 nucleotides (nt) in length - a practical size range - have been introduced into human cells to assess editing efficiency. In this study, we used relatively long sense ssODNs with modifications to investigate the effects of the number and position of modifications on genome editing efficiency in human cells, shedding light on the potential of ssODNs for genome editing.

The modifying nucleic acid used, locked nucleic acid (LNA), is a valuable tool for nucleotide modification; it can be chemically synthesised, and its incorporation can be strategically positioned during the design phase[16,17]. The introduction of LNA into oligonucleotides increases their melting point and induces a C3'-terminal (N-type) sugar structure in the molecular backbone, resulting in a higher binding affinity to DNA and RNA than natural

[1]Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology, 1-8- 31 Midorigaoka, Ikeda 563-8577, Osaka, Japan. [2]Division of Genome Medicine, Institute of Advanced Medical Sciences, Tokushima University, 3-18-18 Kuramotocho, Tokushima 770-8503, Japan. [3]Laboratory of Biofunctional Molecular Medicine, Health and Nutrition, National Institutes of Biomedical Innovation, 7-6-8 Saito-asagi, Ibaraki 567-0085, Osaka, Japan. [4]Department of Environmental Engineering and Green Technology, International Institute of Technology, Universiti Teknologi Malaysia, 54100 Kuala Lumpur, Malaysia. ✉email: maseda.h@aist.go.jp

deoxyribose of the same size[18–20]. Although several examples of genome editing using ssODNs of LNA: DNA chimeras have been reported[12,17,21–24], there are no reports of genome editing of sense strands longer than 60 nt using such chimeric ssODNs. Therefore, in this study, we produced various ssODNs with different numbers and positions of LNA modifications of ssODNs longer than 60 nt, which can now be synthesised with improved technology, and performed genome editing using them to investigate the optimal number and positions of LNAs that significantly contribute to improving the efficiency of genome editing.

Results

Genome editing efficiency based on native DNA length

In order to easily evaluate the possibility of genome editing using ssODNs, human cells (HEK293T) with an inactivated EGFP cassette (Fig. S1) were generated by introducing a cassette gene (Fig. 1a) containing eight foreign bases not required for EGFP. Genome editing was then attempted by transecting cells with 20–100 nt ssODNs (Table 1; Fig. 1a) designed to encode the native GFP sequence, which does not contain the eight bases introduced into the inactivated EGFP. Following transfection, genomic DNA was extracted from the transfected cell population and deep sequenced by NGS to determine the percentage of deletion of the added 8-base sequence. Genome editing efficiencies for 8-base deletions were less than 0.0001%, 0.0002%, 0.0019%, 0.0091% and 0.0154% for 20, 40, 60, 80 and 100 nt, respectively. Thus, although the efficiency was not exceptionally high, site-specific editing could be performed by introducing ssODNs into cells, with the efficiency increasing as the length of the transfected ssODNs increased (Fig. 1b). Flow cytometry analysis of a subset of the transfected cell population also revealed, as expected, an increased number of cells expressing EGFP fluorescence following genome editing as a function of the length of the transfected ssODNs (Fig. 1c). However, the efficiency of this assay was consistently lower than that of NGS. Unlike NGS analysis, which directly examines the sequence of a normally edited genome, flow cytometry evaluates a phenotypic expression system that counts fluorescent proteins normally expressed after genome editing. Therefore, it may not detect cells in an insufficiently expressed state. Nevertheless, flow cytometry analysis is more convenient for this study, in which reproducibility needs to be confirmed in a large number of cells, due to its fast turnaround time and much simpler analysis process

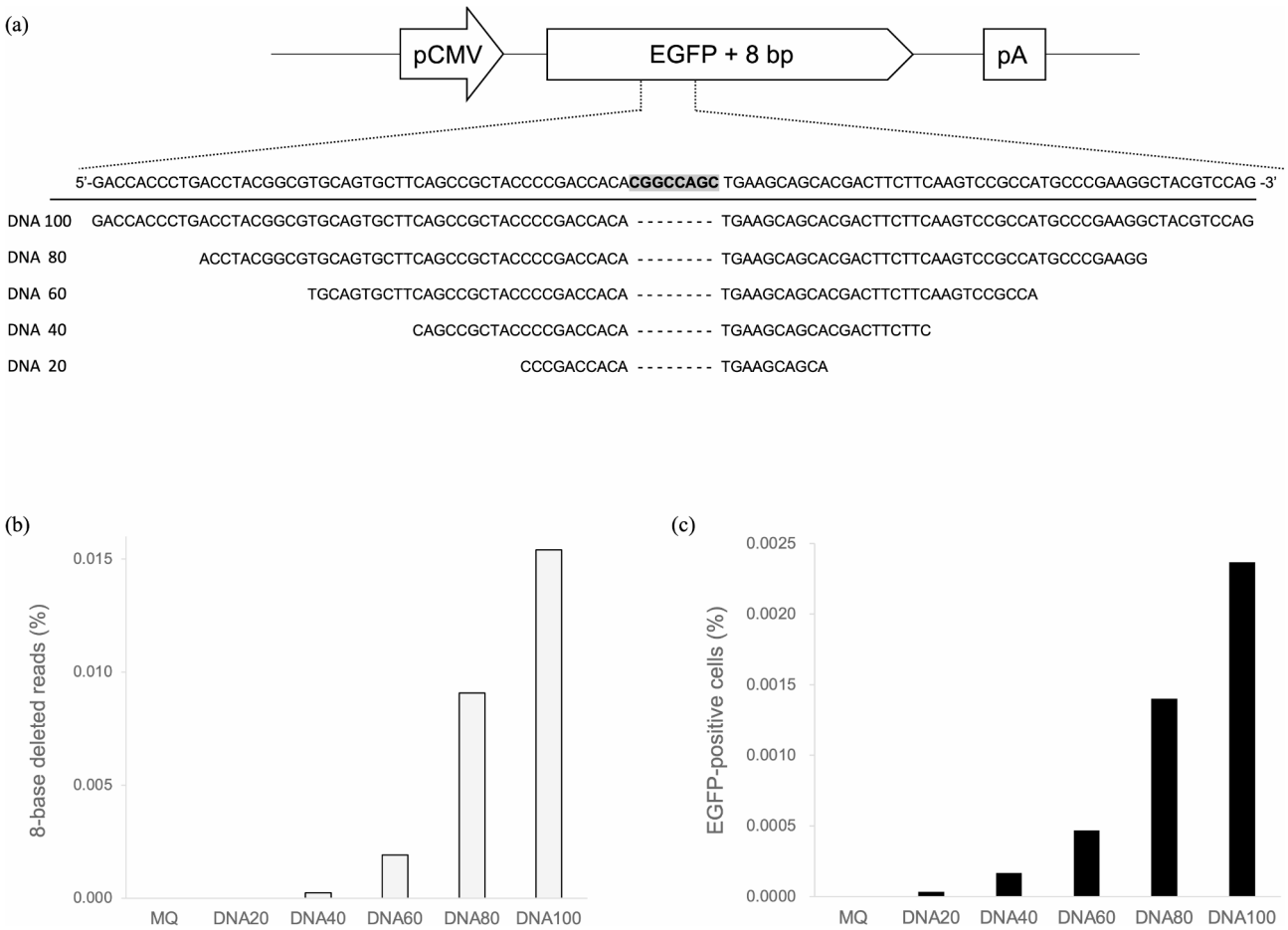

Fig. 1. Verification of genome editing efficiency using native DNAs. (**a**) The sequence of oligonucleotides used for genome editing (**b**) The ratio of sequence reads in which unnecessary 8-base were deleted using a next-generation sequencing analyzer. (**c**) Proportion of EGFP-positive cells using flow cytometry.

Name	Sequence (5' to 3')
DNA-40	CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTC
DNA-50	TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTC
DNA-60	TGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCA
DNA-70	CGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCC
DNA-80	ACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG
DNA-90	CCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACG
DNA-100	GACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG CTACGTCCAG
40–10 L	**CAGCCGCTA**CCCCGACCACATGAAGCAGCAC**GACTTCTTC**
60–10 L	**TGCAGTGCTT**CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCA**AGTCCGCCA**
70–10 L	**CGGCGTGCA**GTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCC**GCCATGCCC**
80–10 L	**ACCTACGGC**GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT**GCCCGAAGG**
90–10 L	**CCCTGACCT**ACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCG**AAGGCTACG**
100–10 L	**GACCACCCT**GACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGG C**TACGTCCAG**
80–12 L	**ACCTACGGCGT**GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCC**ATGCCCGAAGG**
80–14 L	**ACCTACGGCGTGC**AGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCG**CCATGCCCGAAGG**
80–12 L-50c2L	**ACCTACGGCGT**GCA**G**TGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTC**C**GCC**ATGCCCGAAGG**
80–12 L-40c2L	**ACCTACGGCGT**GCAGTGCT**T**CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTC**A**AGTCCGCC**ATGCCCGAAGG**
80–12 L-30c2L	**ACCTACGGCGT**GCAGTGCTTCAGC**C**GCTACCCCGACCACATGAAGCAGCACGAC**T**CTTCAAGTCCGCC**ATGCCCGAAGG**
80–12 L-20c2L	**ACCTACGGCGT**GCAGTGCTTCAGCCGCTA**C**CCCGACCACATGAAGCAGCA**C**GACTTCTTCAAGTCCGCC**ATGCCCGAAGG**
80–12 L-10c2L	**ACCTACGGCGT**GCAGTGCTTCAGCCGCTACCCCG**A**CCACATGAAG**C**AGCACGACTTCTTCAAGTCCGCC**ATGCCCGAAGG**
80–12 L-c2L	**ACCTACGGCGT**GCAGTGCTTCAGCCGCTACCCCGACCAC**AT**GAAGCAGCACGACTTCTTCAAGTCCGCC**ATGCCCGAAGG**
70–10 L-40c2L	**CGGCGTGCA**GTGCT**T**CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTC**A**AGTCC**GCCATGCCC**
70-(10 − 2)L	CG**GCGTGCA**GTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCC**GCCATGCCC**
70-(10 − 2)L-40c2L	CG**GCGTGCA**GTGCT**T**CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTC**A**AGTCC**GCCATGCCC**
90-70type-10 L-40c2L	CCCTGACCTA**CGGCGTGCA**GTGCT**T**CAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTC**A**AGTCC**GCCATGCCC**GAAGGCTACG

Table 1. Single-stranded oligodeoxyribonucleotides used in this study. Native DNAs are shown as typical capital letters (A, C, G, and T) and LNAs shown as bold.

compared to NGS. Therefore, in future studies, we used this method to investigate the conditions of the ssODN method before incorporating NGS analysis for the more detailed studies required (Fig. S2).

These studies confirmed the benefits of increasing ssODN length to improve the efficiency of genome editing using native DNA; however, there is a limit to the length of ssODNs that can be stably synthesised and delivered. In addition, increasing the chain length reduces the efficiency of ssODN incorporation into the nucleus. It is thought that there is a practical limit to the length of the chain that can be increased to significantly improve the efficiency of genome editing. We therefore decided to investigate whether the use of LNA, which is expected to improve binding affinity with genomic DNA, would facilitate genome editing compared to the introduction of unmodified ssODNs, even when the ssODNs have the same strand length.

Genome editing efficiency depending on the number of LNAs introduced

In general, it has been reported that instead of placing LNAs in a contiguous arrangement, inserting an appropriate number of DNAs between LNAs enhances their editing efficiency in genome editing using ssODNs (19, 26). Based on this fact, LNAs were introduced at both ends (3′ and 5′) of 80 nt ssODNs, a length suitable for detecting genome editing efficiency. Three variations were created: ssODNs containing 10 (80–10 L), 12 (80–12 L) and 14 (80–14 L) LNAs to investigate the effect of LNA introduction (Fig. 2, Group 2).

In Fig. 2, the efficiency of unmodified 80nt native ssODN is denoted as 1 to facilitate comparison of length and genome editing efficiency. The LNA results are shown in Fig. 2, Group 2, and the editing efficiency with native ssODNs is also shown in the same graph (Fig. 2, Group 1).

As a result, 80–10 L showed 1.23 times higher efficiency than native DNA of the same length, and 80–12 L with 12 LNAs showed 1.5 times higher efficiency, equivalent to that of 100 nt native DNA. In particular, 80–14 L with 14 LNAs produced more than twice the efficiency of 80–12 L and a 3.77-fold increase in editing efficiency compared to 80-nt native DNA. These results suggest that the introduction of 10 or more LNAs at both ends of an 80-nt ssODN significantly increases the efficiency of genome editing as the number of LNAs introduced increases.

Efficiency of genome editing according to the position of LNAs

LNA introduction was also effective for relatively long ssODNs, with the effect being particularly pronounced in 80–14 L, where an additional pair of LNAs (2 in total) was introduced within 80–12 L (Fig. 2, Group 2). Indeed, the binding affinity and stability of ssODNs vary depending on the number and position of LNAs[14,21,22,25].

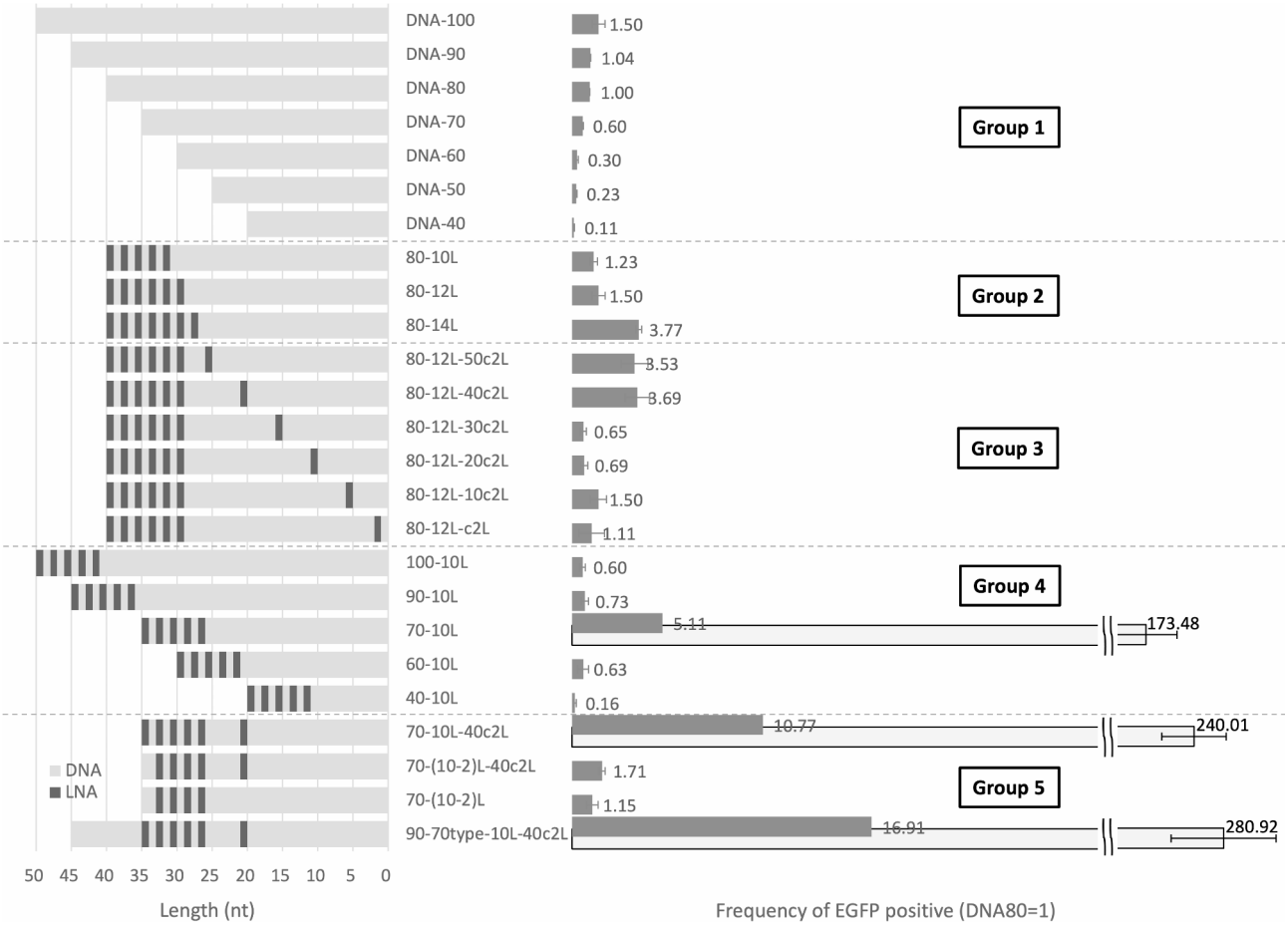

Fig. 2. Genome editing efficiency with different types of ssODNs in 8-base inserted EGFP in HEK293T. The left side of the graph visualises half of the ssODNs used for genome editing. The genome editing target region is located in the centre of the ssODNs, and the nucleic acids exhibit left-right symmetry around the editing region. The graph on the right shows the efficiency of genome editing, with the dark grey bar representing the 8-base deletion efficiency and the light grey bar representing the 1-base substitution efficiency. The efficiency is expressed as a ratio when the 8-base deletion efficiency using native DNA 80nt is set to 1. All assays were performed in triplicate.

However, no data have been reported on the optimal position of 60-nt or longer ssODNs for genome editing in human cells. Therefore, we investigated the effect of LNA modification position on genome editing.

Differences in genome editing efficiency according to LNA position were compared using multiple ssODNs with the innermost LNA changed while maintaining the length and number of LNAs at 80–14 L, which had the highest editing efficiency in previous studies (Fig. 2, Group 3). Interestingly, the results showed significant differences in editing efficiency depending on the location of the additional LNAs, even when the same number of LNAs were introduced. These characteristics could be classified into three patterns, as described below: LNA positions that increase editing efficiency (80–12 L-50c2L and 80–12 L-40c2L), LNA positions that have minimal effect on editing efficiency (80–12 L-10c2L and 80–12 L-c2L), and regions that decrease editing efficiency (80–12 L-30c2L and 80–12 L-20c2L).

In particular, when LNAs were introduced 25 nt from the centre of the ssODN (80–12 L-50c2L) and 20 nt from the centre (80–12 L-40c2L), their efficiency was more than double that of 80–12 L and almost equal to that of 80–14 L, which introduced LNAs at 27 nt from the centre of the ssODN. Genome editing efficiency was significantly increased when LNA was introduced into the region 20–27 nt from the centre of the ssODN (Fig. 2, Group 3).

Subsequently, 80–12 L-10c2L and 80–12 L-c2L, in which LNAs were introduced closer to the centre, showed almost identical editing efficiencies to 80–12 L without any effect of LNA integration. In contrast, 80–12 L-30c2L and 80–12 L-20c2L, in which LNAs were introduced 15 nt and 10 nt from the centre, respectively, showed lower efficiency than native DNA of the same length, despite the 14 LNAs introduced in ssODNs. Thus, the introduction of additional LNAs in inappropriate locations conversely hinders genome editing, suggesting that it is very important to consider the optimal location of LNA modifications (Fig. 2, Group 3).

This observation suggested that the position of LNAs in the ssODN, rather than the absolute number of LNAs introduced, may play an important role in improving genome editing efficiency. Therefore, we further

investigated genome editing efficiency using ssODNs of different lengths with 10 LNAs at the ends (Fig. 2, Group 4).

Interestingly, 90 nt (90–10 L) and 100 nt (100–10 L) sequences showed lower editing efficiencies than native DNA of the same length. Conversely, 60 nt (60–10 L), 70 nt (70–10 L) and 80 nt (80–10 L) showed higher editing efficiencies than their corresponding native DNA sequences of the same length (Fig. 2, Group 4). Among these, a significant increase in editing efficiency was observed for 70–10 L, which was 8.5 times higher than that of native DNA of the same length (DNA-70) and much more effective than conditions involving longer lengths with more LNA introduced (e.g. 80–14 L, 80–12 L-50c2L and 80–12 L-40c2L). In 70–10 L, LNA was introduced 25–35 nt away from the ssODN centre, which partially overlapped with the 20–30 nt region where editing efficiency had previously increased with the introduction of additional LNA. In addition, 60–10 L introduced LNAs at the same position (20–30 nt from the centre of ssODN) where an increase in editing efficiency was observed that was higher than native DNA of the same length (Fig. 2, Group 1, DNA-60). However, in the present study, it was concluded that the 60 nt ssODN was not long enough for genome editing; hence the effect of LNA in this length ssODNs was not pronounced. Therefore, it is clear that in order to increase the efficiency of genome editing methods using ssODNs, it is necessary to use ssODNs of appropriate length, taking into account the position of LNA introduction.

Optimal placement of LNA in ssODNs to improve editing efficiency

In order to investigate the ssODNs that show the most effective genome editing efficiency, we used 70–10 L, which has the highest editing efficiency based on the results of the experiments confirmed so far, as a standard and performed the following studies: (1) addition of LNA within ssODN, (2) deletion of LNA at the end sides of ssODN, (3) adjustment of the length of the nucleic acid (Fig. 2, Group 5).

In the experiment with 80 nt, LNA was added at the same position for 70–10 L because the editing efficiency was improved by adding LNA at a distance of 20–25 nt from the centre of ssODN. As a result, compared to 70–10 L, the editing efficiency of 70–10 L-40c2L was approximately doubled just by adding a pair of LNAs inside each arm (Fig. 2, Group 5). This result shows the highest editing efficiency among the ssODNs to which the identified LNA has been added so far, and the addition of LNA to this region has always been shown to be effective in increasing genome editing efficiency. From these results, it was considered important for efficient genome editing to introduce a sufficient number of LNAs at appropriate intervals at positions 20–35 nt from the centre (ssODN), specifically five pairs at positions 25–35 nt from the centre and also one pair at positions 20–25 nt.

To check the validity of the position and number of LNAs introduced, an ssODN was created by removing one outermost and one innermost LNA from this ssODN. The editing efficiency of 70-(10 – 2)L-40c2L, in which the outermost LNAs were removed from 70 to 10 L-40c2L, decreased dramatically. In addition, the editing efficiency of 70-(10 – 2)L, in which the innermost LNA was removed from 70-(10 – 2)L-40c2L, was lower (Fig. 2, Group 5). In other words, it was again shown that the LNAs in the region 25–35 nt from the centre, which is the position of the 70 nt end, play an important role and maintain high efficiency in cooperation with the LNAs 20 nt from the centre added with 70–10 L-40c2L.

Next, to confirm whether the introduction of the terminal LNA is important for the efficiency of genome editing, as is commonly believed, we added 10 nt native DNA to both ends of 70–10 L-40c2L to create 90 – 70 type-10 L-40c2L, which has an LNA at the same position but is not terminally modified and measured its genome editing efficiency. The results showed that even if the LNA modification position was not changed and the strand length was increased so that there was no LNA at the end, no decrease in efficiency was observed and the genome editing efficiency was 1.5 times higher than that of 70–10 L-40c2L, indicating that the introduction of the terminal LNA is not essential for editing efficiency in the using ssODN method (Fig. 2, Group 5). This result suggests that the presence of LNA at the above position is more important than the presence of LNA at the end of ssODNs.

Effect of LNA-introduced ssODN on base substitution

The introduction of LNAs has succeeded in deleting 8 bases from target genes as intended, with 17.95-fold efficiency (Fig. 2, Group 5, 70-10L-40c2L) compared to native DNA of the same length (Fig. 2, Group 1, DNA-70). In actual therapy, however, it is necessary to replace precisely the gene mutation that causes the genetic disease. Therefore, we decided to investigate the possibility of single nucleotide replacement by ssODNs, using the ssODNs that showed high genome editing efficiency in the previous experiment. The cell used in the experiment, '80stop EGFP cell' (Fig. S3), is a cell in which mutant EGFP has been introduced into the genome by converting the 80th amino acid Lys (AAG) to a stationary codon (TAG). Using the same ssODN as in the previous experiment, it is possible to replace the T with an A instead of a deletion. The results of the efficiency comparison based on DNA-80 are shown in Fig. 2 (Group 4 and 5) and each specific base substitution frequency is shown in Fig. 3. The efficiency of base substitution was approximately 173.48-fold for 70–10 L, 240.01-fold for 70–10 L-40c2L and 280.92-fold for 90-70type-10 L-40c2L when DNA-80 was set to 1. This was 33.95-fold, 22.28-fold and 16.61-fold higher than the efficiency of the corresponding 8-base deletions. As a result of confirming the efficiency of nucleotide sequence substitution using NGS (Fig. 3), which exceeded 0.42% for 70–10 L, 0.52% for 70–10 L-40c2L and 0.63% for 90 – 70 type-10 L-40c2L, it was confirmed that nucleotide substitution was quite efficient. Thus, the efficiency of genome editing with LNA-introduced ssODNs was consistently higher for substitutions than for deletions, demonstrating that genome editing with ssODNs is an excellent genome editing method even for single nucleotide substitutions.

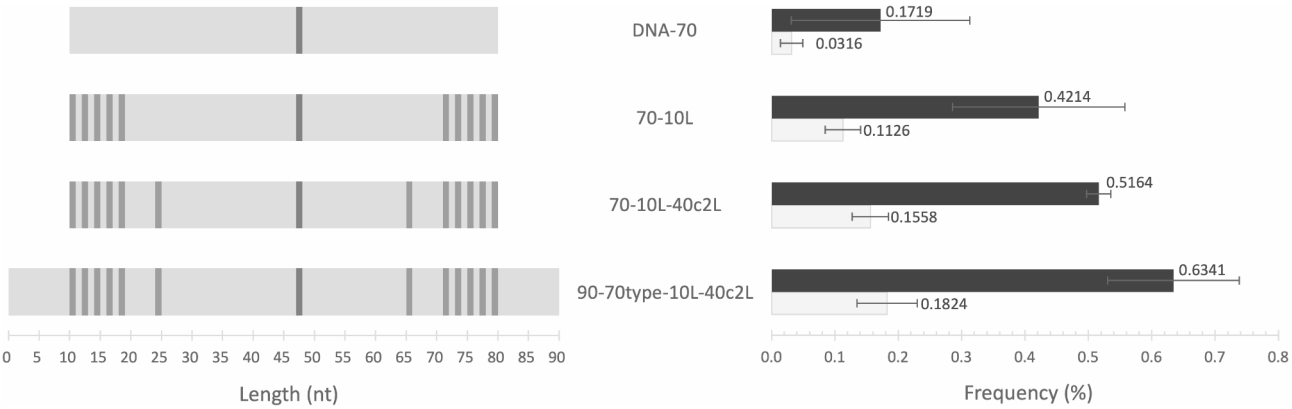

Fig. 3. Investigation of point mutation generation efficiency using ssODNs containing native DNA or introduced LNA. On the left is a schematic diagram of ssODNs inducing base substitution. It has the same base sequence as the ssODNs used to confirm the 8-base deletion efficiency; light grey indicates the DNA and grey indicates the site where LNA was introduced. The dark grey colour in the middle indicates the site of base sequence substitution (T to A). The graph on the right shows the frequency of base sequence conversion when using each ssODN, comparing the results of flow cytometry analysis (white bars) with those of base sequence analysis using NGS (black bars). The flow cytometry results show the ratio of the number of cells expressing EGFP to the total number of counting cells (1,500,000 cells in total), and the NGS analysis results were calculated as the ratio of the number of base sequences in which the target site was substituted from T to A to the number of base sequences that were 100% identical to the 15-base sequences before and after the target base (30 bases in total). All assays were performed in triplicate.

Discussion

Genome editing is becoming an integral part of various industries involving living organisms. In medicine, genome editing promises to be a next-generation therapeutic approach by enabling the artificial repair of disease-causing genes. Although many methods for genome editing have been reported to be highly efficient[26–28], each method has its own advantages and disadvantages when considered for actual therapeutic applications. Among these, the ssODN method does not require the introduction of foreign nuclease protein to edit the genome, and therefore essentially does not require reactions such as forced double-strand breaks during editing, minimising the risk of off-target effects, and if used for genome editing or therapy in individuals, there is no concern about an immune response due to the introduction or expression of foreign proteins. Despite its advantages, the development of genome editing using ssODNs has lagged behind the Nobel Prize-winning CRISPR-Cas method, and further optimisation is needed to address the diversity of genomic medicine.

In ssODN genome editing, it is generally accepted that the endogenous repair system recognises and acts on the D-loop formed at the mismatch site when the ssODN used for genome editing is paired with a complementary sequence in the genome, resulting in editing of the target site[12,14,29]. In this case, the formation of stable D-loops and efficient activation of the endogenous repair system are considered critical for efficient genome editing.

In this study, we aimed to improve the efficiency of genome editing by investigating the optimal length of ssODN and the introduction position of LNA that can more stably bind ssODN and genomic DNA without interfering with repair system factors. The first step was to investigate the optimal length of ssODNs. The results showed that the upward trend in genome editing efficiency increased with ssODN length but decreased sharply when ssODN length increased above 100 nt (data not shown). This decline in efficiency may be partly due to the fact that both the efficiency and the amount of incorporation decrease with increasing molecular size. Next, we attempted efficient genome editing using LNA, an artificial nucleic acid that can increase the binding strength to nucleic acids and their stability without changing their size, despite being a short nucleic acid. As expected, it was confirmed that editing efficiency generally increased as the number of LNAs introduced increased (Fig. 2, group 2).

Interestingly, however, the effect of LNA introduction varied depending on its position. In particular, the efficiency of genome editing decreased significantly when LNAs were introduced at positions more than 40 nt from the centre (90–10 L and 100–10 L, Fig. 2, Group 4) or not more than 10–15 nt from the centre of ssODNs (80–12 L-30c2L and 80–12 L-20c2L, Fig. 2, Group 3). This suggests that the structural characteristics of LNAs should be carefully considered when introducing them into the ssODNs. This finding has important implications for simplifying the use of nucleic acid-only genome editing.

A previous report showed that the introduction of modified nucleic acids at both ends had a significant effect on the efficiency of genome editing with short ssODNs[15,19]. However, our current investigation, using a relatively longer nucleotide length of 70 nt, showed that genome editing efficiency was improved when 10 LNAs were introduced 25–35 nt from the centre of ssODNs rather than at both ends. This positional introduction of LNAs had the most significant contribution to increased efficiency, highlighting the pivotal role of the structural position of the introduced LNA in genome editing efficiency, surpassing the effect of having LNA at both ends. As can be seen from the experimental results, for ssODNs that have a fundamentally stable structure with LNAs at appropriate positions, the introduction of an additional LNA at 20 nt from the centre appears to act

as an enhancer that further improves thermal and structural stability. Thus, strengthening the native DNA by introducing LNA at the appropriate position rather than at the ssODNs terminus would result in more stable binding to the genome and improved editing efficiency.

The optimised ssODNs were also very effective at base substitution. The three most efficient ssODNs in this experiment, 70–10 L, 70–10 L-40c2L and 90-70type-10 L-40c2L, were used as they were and in HEK293-T (sub-GFP) cells, the efficiency of base substitution reached 0.42–0.63%, far exceeding the efficiency of deletion. This efficiency is expected to spur the application of ssODNs in real medical fields, considering that many disease mutations are caused by base substitutions.

Although we have been able to dramatically improve the efficiency of genome editing methods using ssODNs by optimising their length and LNA insertion position, there are still many issues to be resolved before they can be applied to personalised medicine and genome therapy. In particular, the structural characteristics of genomic regions associated with diseases requiring editing are diverse and fluid. In addition to improving the structural safety of ssODNs as described in this paper, much more research is needed to identify the factors involved in genome editing and to address a variety of cases. We are currently continuing our attempts to improve the efficiency of genome editing by identifying and appropriately removing factors that inhibit genome editing with ssODNs and have already achieved favourable results. Combined with the efficiency of genome editing using LNAs established here, we expect to move closer to the realisation of genome editing medicine.

Materials and methods

Cell culture conditions

Enhanced green fluorescent protein (EGFP) reporter cells were cultured in Dulbecco's modified Eagle's medium (Sigma, St. Louis, MO, USA) supplemented with 10% (v/v) fetal bovine serum (Sigma-aldrich, Co., St. Lois, MO, USA) and 1% (v/v) antibiotic/antimycotic solution (Life Technologies, Inc., Grand Island, NY, USA). The cell culture incubator (Thermo Fisher Scientific, Inc.) was maintained at 37˚C with 5% of CO_2 gas concentration.

EGFP reporter cell lines

Modified EGFP with an 8-base insertion was synthesized by Fasmac (Japan) and was amplified by PCR using the inf-EGFP5 (5′-TAGAGCTAGCGAATTATGGTGAGCAAGGGCGAGG AG-3′) and GFP-sv40polyA-pCDH-R (5′-ATTTAAATTCGAATTATAAGATACATTGATGAGTT-3′) primer pair. The resulting product was ligated with Eco RI-treated plasmid pCDH-CMV-MCS-EF1-RFP-Puro (System Biosciences) through the In-Fusion reaction. As this plasmid vector does not have a replication origin that can replicate itself in human cells, it is possible to obtain a strain in which the gene on the vector has been introduced into the genome by selection with a marker gene. The plasmid vector was introduced into HEK293T cells using FuGENE 6 Transfection Reagent (Roche, Basel, Switzerland, now Promega Corporation). After 48 h of transfection, the cells were transferred to a medium containing 0.5 μg/mL puromycin (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) to select for stable clones with integrated EGFP genes, and the cells were cultured for three weeks. After that, more than 20 puromycin-resistant colonies that appeared on the plate were isolated using the limiting dilution method. Srains with functional EGFP were selected by fluorescence observation, while strains with mutated and non-functional EGFP were selected by confirming the genome sequence using the Sanger method, and strains with functional EGFP or mutated and non-functional EGFP in the genome were selected. Finally, strains with stable phenotypes were selected as cells for genome editing analysis.

Oligonucleotides

ssODNs and LNAs were designed as sense strands of genomic DNA those sequence information is in the Table 1 ssODNs were purchased from Eurofins (Tokyo, Japan) and Macrogen (Seoul, Korea). All LNA oligonucleotides used in this study were purchased from Ajinomoto Bio-Pharma (Osaka, Japan) and were HPLC-purified to a manufacturer-guaranteed grade.

Transfection of ssODN

Initially, 2.5×10^5 cells were seeded in 3 ml DMEM (contained 10% FBS and 1% antibiotic/antimycotic) of a 6-well plate and then cultured for 48 h to reach approximately 80% confluence. All types of ssODNs (5.5 μg) were diluted with OptiMEM (Life technoligies, NY, USA) and FuGENE® HD transfection Reagent (10 μl, Promega, Madison, WI, USA) was added to achieve the final volume of 165 μl. The mixture was incubated for 5 min at room temperature (24 ± 2 °C), and then 150 μl of mixture was added to each well of a 6-well plate containing 3 ml of cells in the growth medium. The plate was returned to the incubator for 24 h. The supernatant was removed, and the cells were transferred to a 100 mm dish (Eppendorf, Hamburg, Germany) and continuously cultured for 4 days until the cells reached confluence (reach approximately 90 ~ 100%).

Flow cytometry assay

EGFP-positive cells were quantified using an Attune NxT Acoustic Focusing Cytometer (Thermo Fisher Scientific, Inc.) and analyzed with Attune NxT software version 2.7.0. Confluent cells were harvested and resuspended in 5 ml cell culture medium. The analysis rate was 200 μl/min, and the counting continued until the total cell number reached 1.5×10^6 cells.

Genomic DNA sequencing

Genomic DNA was extracted from approximately 2×10^6 cells harvested cells by cytometry measurements using NucleoSpin tissue (MACHEREY-NAGEL, Germany) and subsequently PCR-amplified for 30 cycles using primeSTAR Max DNA polymerase (Takara, Japan) with tailing primers. To confirm the results from sufficient number of genomes, 1600 ng of genome corresponding to 2×10^5 cells was used as a template for

PCR. The resulting amplicons were purified using a NucleoSpin Gel and PCR Clean-up kit (MACHEREY-NAGEL, Germany). The samples were further subjected to 10 cycles of PCR using the same DNA polymerase with TruSeq CD index (Illumina, USA) for indexing, followed by purification with AMPure (Beckman, USA). Purified DNA samples were then quantified with a QubitTM 4.0 fluorometer (Invitrogen by Thermo Fisher Scientific, Singapore), pooled in equimolar ratios, and analyzed via high-throughput DNA sequencing at the target loci. The sequencing was carried out using the iSeq100 (Illumina, USA), and the results were analyzed using CLC Genomic Workbench 20 (QIAGEN Digital Insights, Denmark). As a result of sequencing, the number of base sequences that were 100% identical to a total of 38-base sequences, 15-bases before and after the 8-base constituting the target loci, was counted. In addition, in the case of point mutations, a total of 31-base sequences, including one base at the target point and 15-base before and after, were counted using the same method. Genome editing efficiency was expressed as the ratio of the read numbers showing the 8-base deletion (or point mutation) to the sum of the read numbers with no change in the sequence and the read numbers indicating 8-base deletion (or point mutation).

Data availability

The raw sequencing data have been deposited in the DDBJ database under accession number PRJDB19145 and are available for public access.

Received: 3 December 2024; Accepted: 11 March 2025
Published online: 21 March 2025

References

1. Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
2. Jinek, M. et al. A programmable Dual-RNA–Guided DNA endonuclease in adaptive bacterial immunity. *Sci. (80-).* **337**, 816–821 (2012).
3. Cho, S. W., Kim, S., Kim, J. M. & Kim, J. S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).
4. Zhang, X. H., Tee, L. Y., Wang, X. G., Huang, Q. S. & Yang, S. H. Off-target effects in CRISPR/Cas9-mediated genome engineering. *Mol. Ther. - Nucleic Acids*. **4**, e264 (2015).
5. Yang, Y., Xu, J., Ge, S. & Lai, L. CRISPR/Cas: advances, limitations, and applications for precision cancer research. *Front. Med.* **8**, 649896 (2021).
6. Fu, Y. et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
7. Charlesworth, C. T. et al. Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* **25**, 249–254 (2019).
8. Richardson, P. D., Kren, B. T. & Steer, C. J. Targeted gene correction strategies. *Curr. Opin. Mol. Ther.* **3**, 327–337 (2001).
9. Copeland, N. G., Jenkins, N. A. & Court, D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nat. Rev. Genet.* **2**, 769–779 (2001).
10. Aarts, M. & te Riele, H. Subtle gene modification in mouse ES cells: evidence for incorporation of unmodified oligonucleotides without induction of DNA damage. *Nucleic Acids Res.* **38**, 6956–6967 (2010).
11. Komor, A. C., Badran, A. H. & Liu, D. R. Editing the genome without double-stranded DNA breaks. *ACS Chem. Biol.* **13**, 383–388 (2018).
12. van Ravesteyn, T. W., Dols, A., Pieters, M., Dekker, W., te Riele, H. & M. & Extensive trimming of short single-stranded DNA oligonucleotides during replication-coupled gene editing in mammalian cells. *PLOS Genet.* **16**, e1009041 (2020).
13. Radecke, S., Radecke, F., Peter, I. & Schwarz, K. Physical incorporation of a single-stranded oligodeoxynucleotide during targeted repair of a human chromosomal locus. *J. Gene Med.* **8**, 217–228 (2006).
14. Igoucheva, O., Alexeev, V. & Yoon, K. Targeted gene correction by small single-stranded oligonucleotides in mammalian cells. *Gene Ther.* **8**, 391–399 (2001).
15. Wu, X. S. et al. Increased efficiency of oligonucleotide-mediated gene repair through slowing replication fork progression. *Proc. Natl. Acad. Sci.* **102**, 2508–2513 (2005).
16. Parekh-Olmedo, H., Drury, M. & Kmiec, E. B. Targeted nucleotide exchange in Saccharomyces cerevisiae directed by short oligonucleotides containing locked nucleic acids. *Chem. Biol.* **9**, 1073–1084 (2002).
17. Petersen, M. & Wengel, J. LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol.* **21**, 74–81 (2003).
18. Liczner, C., Duke, K., Juneau, G., Egli, M. & Wilds, C. J. Beyond ribose and phosphate: selected nucleic acid modifications for structure–function investigations and therapeutic applications. *Beilstein J. Org. Chem.* **17**, 908–931 (2021).
19. Crinelli, R. Design and characterization of decoy oligonucleotides containing locked nucleic acids. *Nucleic Acids Res.* **30**, 2435–2443 (2002).
20. Owczarzy, R., You, Y., Groth, C. L. & Tataurov, A. V. Stability and mismatch discrimination of locked nucleic Acid–DNA duplexes. *Biochemistry* **50**, 9352–9367 (2011).
21. Sun, B. W. et al. Sequence and pH effects of LNA-containing triple Helix-Forming Oligonucleotides: physical chemistry, biochemistry, and modeling studies. *Biochemistry* **43**, 4160–4169 (2004).
22. Georgiadou, M. et al. Intramuscular evaluation of chimeric locked nucleic Acid/2′OMethyl-Modified antisense oligonucleotides for targeted exon 23 skipping in Mdx mice. *Pharmaceuticals* **14**, 1113 (2021).
23. Crooke, S. T., Liang, X. H., Baker, B. F. & Crooke, R. M. Antisense technology: A review. *J. Biol. Chem.* **296**, 100416 (2021).
24. Andrieu-Soler, C. Stable transmission of targeted gene modification using single-stranded oligonucleotides with flanking LNAs. *Nucleic Acids Res.* **33**, 3733–3742 (2005).
25. Vester, B. & Wengel, J. L. N. A. Locked nucleic acid: High-affinity targeting of complementary RNA and DNA. *Biochemistry* **43**, 13233–13241 (2004).
26. Zhang, L. et al. Boosting genome editing efficiency in human cells and plants with novel LbCas12a variants. *Genome Biol.* **24**, 102 (2023).
27. Zhang, Z. et al. Efficient engineering of human and mouse primary cells using peptide-assisted genome editing. *Nat. Biotechnol.* **42**, 305–315 (2024).
28. Anzalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
29. Sansbury, B. M. & Kmiec, E. B. On the origins of homology directed repair in mammalian cells. *Int. J. Mol. Sci.* **22**, 3348 (2021).

Acknowledgements
This research was supported by AMED under Grant Number JP20ck0106410, and JSPS Grant-in-Aid for Challenging Research (Pioneering) Grant Number JP20K20640. The authors would like to thank Satomi Masai for technical assistance with the study. In addition, we would like to thank BioRENDER for graphic image of this research and Editage (www.editage.com) for English language editing, respectively.

Author contributions
S.K. carried out the formal analysis, investigation, methodology and visualisation, and contributed to the drafting of the original manuscript and its revision and editing. Y.M. and T.K. provided essential resources, including the genetically modified cell lines used in the experiments. H.M. contributed to the conception, formal analysis, investigation, acquisition of funding, supervision and drafting of the original manuscript, and revision and editing. All authors have reviewed the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-94071-5.

Correspondence and requests for materials should be addressed to H.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.