INTRODUCTION

Immunity to infection depends on adequate intracellular signalling in many different immune cell types. Elucidation of these signalling pathways and the function of single signalling proteins have often been performed in cell lines and rodents using a variety of protein-specific inhibitors, knock-down assays and knockout techniques to target the protein of interest. Although these techniques have proven effective, many signalling molecules contain multiple functional domains. Thus, knocking out an entire gene may hide domain-specific activity of the encoded protein. To explore the function of specific subdomains, cDNA encoding recombinant proteins with domain-specific mutations can be forcibly expressed in both primary cells and in cell lines. However, such exogenous expression of cDNA may not be regulated by cellular processes and the resulting recombinant protein might compete with the endogenous protein.

Methods for genome editing of specific loci in intact cells in order to alter endogenous protein activity through specific
single point mutations have been hampered by low efficiency and technical difficulties. However, with the recent discovery of clustered regularly interspaced palindromic repeats (CRISPR)/Cas (CRISPR-associated protein) system, a prokaryotic adaptive immune system, researchers now have a powerful gene-editing tool to precisely introduce gene edits in various living cells and organisms.

In general, the molecular machinery of CRISPR/Cas consists of two major components: a DNA binding domain directing the machinery to a defined locus and an effector domain that facilitates DNA cleavage resulting in a double-strand break (DSB). While several CRISPR/Cas systems exist, the most widely used system adapted for genome editing is CRISPR/Cas9, derived from *Streptococcus pyogenes*. Cas9 is an endonuclease, evolved to be guided to the target site by a single-guide RNA (sgRNA) that binds to a complementary target DNA strand. Specific cleavage of a double-stranded DNA by Cas9 requires binding of Cas9 to the sgRNA-target DNA complex as well as the presence of a protospacer adjacent motif (PAM) in the target DNA sequence. In this case, the Cas9 recognizes a 5′-NGG PAM sequence immediately downstream of the sgRNA binding site (Figure S1).

In a eukaryotic cell, there are several DNA mechanisms in place to repair the DSB. The most common and rapid repair occurs through non-homologous end joining (NHEJ), where micro-homologous overhangs are used to join the DSB. Although small insertions or deletions may occur in NHEJ, it is relatively accurate. In contrast, microhomology-mediated end joining, another DSB repair mechanism, aligns broken strands with homologous sequences of 5-25 bp in length and is highly error-prone. Microhomology-mediated end joining often results in insertions, deletions (indels) and inversions; some more frequently occurring than others. A majority of indels result in frameshifts of the sequence and consequently occurrence of premature stop codons, resulting in knockout mutations. The CRISPR/Cas9 system works with high efficiency, and as such has become a popular tool to create knockout mutations.

Alternatively, DNA breaks can be repaired by homology-directed repair (HDR). By combining a CRISPR/Cas9 cleavage with DNA templates, defined point mutations can be introduced. The DNA template for HDR can be provided as a single-stranded deoxyribonucleotide (ssODN), commonly referred to as a repair template, and contains homologous arms complementary to either the target or non-target DNA strand. Recently, it has been shown that knock-in mutations work with higher efficiency when ssODNs are complementary to the non-target DNA strand, as this is more available to binding post-cleavage. The repair template ssODNs contain the desired knock-in mutation and an additional synonymous knock-in mutation for the PAM site to prevent repeated binding of sgRNA-Cas9 and re-cutting of the target strand. Together with the host DNA repair machinery, the ssODN mediates HDR which introduces the desired mutations into the gene of interest (Figure S1C,D).

To generate mutated cell lines enabling better characterization of T cell signalling pathways, we have employed the CRISPR/Cas9 system to introduce targeted point mutations by HDR in Jurkat T cells using ssODNs. Specifically, we are interested in how phosphorylation associated domains of lymphocyte-specific protein tyrosine kinase (LCK) and T cell-specific adaptor protein (TSAd) function in T cell activation. We thus sought to introduce knock-in mutations in the *LCK* and *SH2D2A* genes that encode for LCK and TSAd respectively, in order to either abolish phosphorysine binding (TSAd-SH2) or to mimic phosphorylation-dependent alteration of SH2 domain specificity (LCK-SH2). Additionally, we mutated two other phosphotyrosines: LCK Y505 and TSAd Y290, which both are known ligands for the LCK SH2 domain. In parallel efforts, we performed experiments to generate Jurkat T cells homozygous for the non-synonymous single nucleotide polymorphism (SNP), rs926103, in *SH2D2A*. This SNP is known to affect the interaction of TSAd with LCK.

Altogether we attempted to edit five loci in these two genes in multiple knock-in experiments. We successfully introduced knock-in mutations within exon 6 and exon 13 of the *LCK* gene and in exon 7 of *SH2D2A*. However, knock-in mutation experiments in the early exons of *SH2D2A* were unsuccessful, implying that the gene-editing system may be influenced by its target site.

We successfully generated knock-in mutated Jurkat cells in repeated independent experiments for three of the five targeted loci. There were striking differences in mutation efficiency within the same gene. Based on our results, we here provide an overview of a simple workflow used for generating and identifying knock-in mutations in cell lines, and we discuss potential challenges faced when editing constitutively expressed genes compared to genes that are upregulated following T cell activation.

## MATERIALS AND METHODS

### 2.1 Reagents

The plasmids used were pX330-U6-Chimeric_BB-CBh-hSpCas9, pSpCas9(BB)-2A-GFP (PX458) (Addgene), pEGFP-N1 (Clontech), pEF-HA-TSAd (encoding S52). sgRNA encoding oligonucleotides targeting the loci of interest were cloned into the pX330 or PX458 vectors. pEF-HA-TSAd-N52 was obtained using Quick Change Mutagenesis (Thermo Fischer Scientific). All new constructs were verified by sequencing. Oligonucleotides for generating sgRNA encoding plasmids, ultramer ssODN for HDR repair, as well as ssODN for PCR and sequencing referred to in the text were booster DOI.
custom synthesized by Integrated DNA Technologies. Primary antibodies used were: anti-CD3ε (clone OKT-3; American Type Culture Collection), anti-LCK (unconjugated or conjugated to AF647, clone 3A5, Santa Cruz Biotechnology), anti-SH2D2A (DyLight 488, clone 3C7; Origene), anti-phosphotyrosine (clone 4G10, Upstate Biotechnology), anti-GADPH (clone 6C5; Chemicon) and anti-TSAd antibodies raised against synthetic peptides of TSAd or the recombinant C-terminus of TSAd (a generous gift from Virginia Shapiro). Secondary antibodies used were: AF647 goat anti-mouse IgG (H+L) (Thermo Fisher), HRP-conjugated goat anti-mouse (H+L) and HRP-conjugated goat anti-rabbit antibodies (Jackson ImmunoResearch).

2.2 | Cell cultures

Jurkat TAg cells, expressing large T antigen (Jurkat T cells) were cultured in complete RPMI 1640 media (supplemented with 10% foetal calf serum, 1 mmol/L sodium pyruvate, 10 mmol/L HEPES, 1 mmol/L MEM nonessential amino acids, penicillin [100 U/mL], streptomycin [100 µg/mL] [all from GIBCOBRL®, Thermo Fischer Scientific] and 50 µmol/L B-mercaptoethanol [Sigma]). To reduce the impact of clonal heterogeneity in the gene-editing experiments, Jurkat T cells were subcloned by limiting dilution. Clones were assessed for expression of LCK, TSAd and surface CD3 by Western blotting and flow cytometry. Clones which displayed similar amounts as the median values of bulk Jurkat T cells were selected for the use in subsequent experiments.

2.3 | Cell stimulation

For activation of Jurkat cells by OKT3, a T75 flask was coated with 10 mL of 5 µg/mL OKT3 for 1 hour. 1 x 10^7 cells were then activated at 37°C, 5% CO₂ for 16hrs. For activation by PMA/ionomycin, 1 x 10^7 Jurkat cells were incubated with 50 ng/mL PMA and 500 ng/mL ionomycin at 37°C, 5% CO₂ for 16 hours. After activation, cells were washed with PBS and resuspended in complete RPMI prior to transfection experiments. For short-term anti-CD3 stimulation, up to 10^7 cells/mL in PBS were prewarmed at 37°C for 5 minutes. Cells were stimulated with 5 µg/mL of anti-TCR antibody (OKT3) for 2 minutes. Stimulation was stopped by adding 1 mL of cold PBS and cells were collected for lysis.

2.4 | Design of sgRNAs and mutation donor ssODNs

PAM sites were identified both on the sense and the non-sense strand in the immediate vicinity of the loci to mutate. Potential sgRNA (20-23 bp) located 5’ to these PAM sites were identified and evaluated using CRISPOR. From one to three sgRNA sequences for each locus were chosen for experiments. 103-127 nucleotides (nt) long ssODNs complementary to the non-target strand were designed based on the genomic sequence centred on the locus to mutate. Some of the ssODNs were symmetric with the Cas9 cut site located in the middle of the ssODN, while later in the course of experiments asymmetric ssODNs, with the Cas9 cut site approximately 36 nt away from the cut site on the 5’-end of the non-target strand, were generated. All ssODNs contained the desired codon and mutations of the PAM site when necessary.

2.5 | Cell transfection

1 x 10^7 Jurkat T cells were transfected with 0.5µg of pEGFP-N1 plasmid, 5 µg of sgRNA-pX330 plasmid as well as 1 µmol/L of donor ssODN (IDT) (containing the desired mutation at the targeted locus as well as the adjacent PAM site) at 240V for 25 ms using a BTX electroporator (Genetronix). Three days after transfection, a limiting dilution in 96-well plates with 1 cell per well was performed. After two-three weeks of expansion, clones were screened for the defined knock-in mutation by allele-specific PCR. Successful knock-ins were assessed for CD3 expression as well as for the protein of interest using flow cytometry and Western blotting.

2.6 | Genomic DNA extraction

To obtain genomic DNA for genotyping, Jurkat T cells were lysed in Tris-EDTA buffer (pH 8.0), 0.1% SDS (BioRad) and incubated with 0.2 µg/µL proteinase K (Sigma) at 55°C for 3 hours and subsequently at 80°C for 10 minutes to deactivate proteinase K. Cell lysates were diluted 20 times with purified water before being used for genotyping. DNA from nucleated blood cells of healthy anonymous donors from the Oslo University hospital blood bank was isolated using rapid salting-out method.

2.7 | Genotyping by allele-specific PCR and sequencing

Allele-specific PCR (AS-PCR) assays were used to identify knock-in mutations. Wild-type (WT) and knock-in–specific primers differed in the 3’ end of the primers by two or more nt, representing the WT or mutated codon sequence or the PAM site mutation. In most cases, to increase the specificity of the primer, a mismatch was introduced 3-4 nt away from the 3’ end of the primer. To increase the specificity of
the PCR reaction, in some experiments betaine (final concentration 1.3 mol/L) was added. The PCR protocol with Taq Polymerase (PROMEGA) used for all AS-PCRs was: 95°C for 2 minutes, 40 cycles for 95°C for 15 seconds, 58°C for 15 seconds, 72°C for 15 seconds and final extension at 72°C for 7 minutes. Resulting PCR products were analysed by electrophoresis at 90V for 30min on a 2% agarose gel. Samples for Sanger sequencing, to confirm knock-in mutations, were prepared by using primers flanking the knock-in region and the same PCR protocol as described above. Amplicons were purified with Wizard SV Gel and PCR Clean-Up System (PROMEGA) and custom sequenced (GATC Biotech).

2.8 | Sequence analysis of heterozygous mutants

The sequencing results were translated by the DNA Baser Assembler (Heracle BioSoft) to the standard nt ambiguity codes and corrected manually, if necessary. The resulting sequences were aligned to the WT sequence of the targeted gene using pairwise sequence alignment with EMBOSS Needle. Heterozygous sequences were deconvoluted by manipulating the parameters of the algorithm and by manual rearrangements.

2.9 | Flow cytometry

Cells were barcoded with three different concentrations of CellTrace Violet (ThermoFisher) according to the manufacturer’s protocols. Different clones were either kept unlabelled or stained with the varying concentrations of CTV and combined respectively. Surface stain (CD3): The combined samples were incubated for 20 minutes on ice with anti-CD3ε (OKT3) and fixed with 2% paraformaldehyde. Subsequently, fixed cells were incubated with AF647 goat anti-mouse IgG (H+L) antibodies. Intracellular stain (TSAd, LCK): The combined samples were fixed with 2% paraformaldehyde at room temperature, permeabilized with 0.1% saponin for 15 minutes and subsequently stained for 1h at room temperature with antibodies: anti-LCK (AF647), anti-SH2D2A (DyLight 488). Barcoded, fluorescently labelled samples were analysed with FACSCanto II flow cytometer (BD Biosciences) and FlowJo software (Tree Star, Ashland).

2.10 | Lysis and Western blotting

Stimulated or treated cells were pelleted and lysed with 0.1% LDS/1% Triton solution containing 0.5% Triton X-100 (VWR International SAS), 50 mmol/L HEPES (GIBCOBRL®, ThermoFisher), 0.05% LDS (Merck), 0.05 mol/L LiCl, 0.5 mmol/L PMSF, 2.5 mmol/L EDTA (pH 8.0), 1 mmol/L sodium vanadate and 1x SIGMAFAST protease inhibitor cocktail (all from Sigma). Cells were lysed for 45 minutes on ice. Lysates were sonicated briefly to break the DNA. The appropriate volume of loading buffer (containing 0.35 mol/L Tris HCl, 10% SDS, 6% B-mercaptoethanol (all Sigma), 30% glycerol (VWR international SAS), 0.175 mmol/L bromophenol blue (Fluka Ag), pH 6.8) was added to the samples. All samples were denatured by boiling at 95°C for 10 minutes. Denatured samples were run with SDS-PAGE and transferred onto PVDF membranes (Bio-Rad Laboratories) with Trans-Blot Turbo Transfer System (Bio-Rad laboratories). The membranes were blocked for 1h with 3% skimmed milk or, when probed with anti-pTyr antibodies, 3% bovine serum albumin (Bio-Rad laboratories) in tris-buffered saline (pH 7.4) with 0.1% Tween-20 (Sigma-Aldrich). The membranes were subsequently incubated overnight at 4°C with primary antibodies diluted in respective blocking buffers. On the next day, the membranes were incubated with the appropriate HRP-conjugated secondary antibody for 1h at room temperature. SuperSignal West Pico Stable Peroxide Solution (Pierce) was used to visualize bands using ChemiDocTM Imaging System (Bio-Rad Laboratories).

3 | RESULTS

3.1 | Heterogeneity of cell lines may affect interpretation of genome editing results

In this work, we aimed to create point mutations in LCK (i.e., Y192F, Y192E and Y505F) (Figure S2A), and in SH2D2A (i.e., N52S, R120K, Y290F and Y290E) (Figure S2B). Both genes encode proteins mediating phosphorytrosine dependent signalling in T cells. Here we report our results from a total of 31 gene-editing experiments of LCK and SH2D2A in Jurkat cells performed over a period of three years. Our current overall mutation and cloning workflow, based on results presented in the following section, is depicted in Figure 1.

Similar to other cancer cell lines, Jurkat T cell lines are inherently heterogeneous.\(^{22}\) During the course of our studies, we noticed that our Jurkat T cells had heterogeneous expression of surface CD3, which may be a clonotypic property. Our proteins of interest, LCK and TSAd, directly or indirectly interact with CD3 to initiate or modulate downstream signalling events in the T cell. Therefore, it was essential that Jurkat T cell clones with uniform expression of surface CD3, LCK and TSAd were selected for our gene-editing experiments. Thus, prior to performing genome editing experiments involving isolation of single clones by limiting dilution, we subcloned our Jurkat T cell line creating monoclonal founder cells, hereafter referred to as WT. The resulting clones were analysed for response to anti-CD3 stimulation, expression of...
surface CD3 and expression of our two proteins of interest, LCK and TSAd (Figure 2). The results showed considerable variability in the Jurkat T cell clones’ response to anti-CD3 stimulation when assessed with anti-phosphotyrosine immunoblotting (Figure 2A). Some clones (including clone 6) responded more strongly than the mother cell line (Jurkat), while other clones (including clone 7) responded relatively weaker. This variability was better correlated with the amount of CD3 expressed on the cell surface than to the amount of expressed LCK (Figure 2C).

3.2 Establishment of method for reliable screening of knock-in mutants

To identify successful targeted point mutations in the Jurkat T cell clones as well as in bulk cells prior to limiting dilution, we used allele-specific PCR (AS-PCR) (Figure 1), a technique also used to detect SNPs.\textsuperscript{23,24} For all loci, we used a common primer in one orientation and two alternate detection primers in the other orientation. The detection primers were complimentary at the 3′-end either to the WT sequence or to the knock-in mutation, including the PAM site mutation if relevant (Figure S3).

Despite lacking exonuclease activity, Taq polymerase tolerates a certain degree of 3′ mismatch in the primer.\textsuperscript{25} To further destabilize the primer-template complex and ensure specific amplification of either WT or the mutated variant, we in addition introduced mismatches 2-5 bases upstream of the 3′-end in the detection primers.\textsuperscript{26} The effect of one such additional mismatch in the +3 position from the 3′ end was systematically tested for the rs926103 SNP (with alleles here referred to as N52 [asparagine] and S52 [serine]) (Figure 2D). As seen in Figure 2E, when the detection primers lacking a destabilizing +3 nt were tested against cDNA-containing plasmids encoding one or the other of the two variants, none of them reliably distinguished the two variants from each other. However, for both detection
primers, inclusion of an A base in position +3 was sufficient to discriminate one allele from the alternative variant. The two alternate bases C or G, had a similar effect only for the primer detecting the allele encoding N52 (Figure 2E).

As genomic DNA is more complex than a cDNA-containing plasmid, we subsequently used the TSAd-N/S52 detection primers containing the +3A nt to genotype the DNA of four anonymous blood donors. Donors 1, 2 and 4 were found to carry the S52 allele, while donor 3 was found to carry the N52 allele. Addition of betaine to the PCR mixture, clearly improved the specificity of the PCR, and revealed that donor 3 also carried the S52 allele (Figure 2F). These genotyping results were confirmed by sequencing (Figure 2G). Taken together, establishment of AS-PCR for testing of knock-in mutations should be done prior to gene-editing experiments. Systematic testing of alternate mismatches in the 3′-end of detection primers for AS-PCR should be performed, to ensure a specific screening assay for knock-in mutations. This is particularly important when the desired knock-in mutation only differs by one or two nt from the WT sequence.

3.3 Transfection and bulk analysis of sgRNA targeting efficiency

Co-expression with cDNA encoding green fluorescent protein (GFP) may allow monitoring of transfection efficiency with plasmids encoding both sgRNA and Cas9 (Figure 3A). Initial experiments using the alternative PX458 plasmid encoding both sgRNA, Cas9 and GFP in cis, showed very low GFP expression efficiency (<10% of positive cells) (Figure 3B). Attempts to isolate GFP positive cells by cell sorting did not yield in any mutated cells (data not shown). In all experiments reported here, we thus co-transfected cells with px330 plasmids encoding the sgRNA and Cas9 together with a plasmid encoding GFP (pEGFP-N1) (Figure 3B). As the GFP was expressed by 30%-50% of cells, we did not consider it necessary to enrich for the GFP+ population with cell sorting.

Limiting dilution to obtain clonal cells is time-consuming, thus if possible, a few days post-transfection, the bulk of cells should be monitored for CRISPR-Cas9 targeting.
efficiency prior to further experimental work. With AS-PCR established prior to genome editing (Figure 2E-G), the bulk of transfected cells can be tested for the presence of knock-in mutations already at day 3 post-transfection, as shown in Figure 3C. However, we noticed that when the detection primers were complementary to the ssODN repair template, false-positive AS-PCR may result, due to excess ssODN present in the media or intracellularly after the transfection (data not shown). In addition, provided the targeted gene product can be detected by flow cytometry, initial experiments with sgRNA-Cas9 transfection in the absence of ssODN repair template may allow assessment of Cas9 targeting efficiency through loss of the target protein in the bulk of transfected cells, as a consequence of generated knockouts. An example of this in the bulk cells expressing Cas9 and sgRNA targeting Lck Y192 is shown in Figure 3D,E.

3.4 | Controlling the heterogeneity of generated mutants

To obtain single clones with homozygous knock-in mutations, the cells were subjected to limiting dilution three days after the transfection. Clones were subsequently screened for successful knock-in mutations 2-3 weeks later, using AS-PCR (Figure 4A). Positive clones were further tested with detection primers for the WT allele (Figure 4B), to reveal possible heterozygosity. Further characterization of the clones included analysis of expression of the targeted protein by Western blotting (Figure 4C,D) and flow cytometry (Figure 4E). Western blotting may reveal changes in protein expression or length. Thus, performing a high percentage PAGE (12% or 4%-20% gradient gels depending on the protein of interest) could be a measure of precaution. One of the LCK Y192F mutants in Figure 4C expresses an additional variant of the LCK molecule of double size (the expected size of LCK is 56kDa). The same blot exposed for a longer time revealed that some of the LCK KO mutants express truncated variants of various lengths of the LCK molecule (Figure S4). Such changes cannot be detected by flow cytometry. Western blot analysis is especially useful for target sites at the C-terminus (eg Lck Y505), as they are unlikely to generate a knockout, but the change in protein length can be observed. Presence of more than one clone in a sample may be detected by flow cytometry, provided that one of the clones is hemizygous or a knockout for the protein of interest (Figure 4E). Since both of our target genes affect T cell activation and TCR-signalling, we also monitored surface expression of the CD3 again. As shown in Figure 4F, in the LCK Y192F mutants surface CD3 expression varied considerably between the clones, depending on whether the mutants were derived from founder clone 1 or 2. In contrast, the LCK KO clones that were isolated as by-products of our mutation procedure, showed a trend towards lower surface expression of CD3, independent of the founder cell line used to generate the mutants. This result...
underlines the importance of performing CRISPR mutagenesis in multiple founder clones, to reduce the influence of clonal heterogeneity on the phenotype of the mutants.

### 3.5 Sequencing of knock-in mutants reveal a wide range of mutations and deletions

The screening methods described above provide most of the information needed to identify knock-in mutants. However, these screening methods may not reveal knockout mutants resulting from indels at the C-terminus, nor will they reveal possible additional silent mutations or mutations resulting in amino acid substitutions outside of the targeted codon. Thus, all clones must also be verified by sequencing.

During the course of our experiments, we sequenced amplicons of approximately 250 bp covering the targeted locus in a total of 54 clones for LCK Y192, 10 clones for LCK Y505, 15 clones for TSAd Y290 and 2 clones for TSAd R120. For LCK Y192 and TSAd Y290, clones with suspected gene disruption were also included for sequencing, explaining the higher number of sequenced clones in comparison with the other loci.

**FIGURE 4** Screening knock-in clones with AS-PCR, Western blotting and flow cytometry. (A) 2-3 weeks after limiting dilution, single clones were screened by AS-PCR to identify clones harbouring the knock-in mutation. Initial screening was performed with a 96-well 2% agarose gel, here testing the presence of Lck Y192F mutation. Samples with positive AS-PCR result are shown with white boxes. (B) Positive hits from the initial 96-well screening were verified with a better resolution 16-well 2% agarose gel, as well as with additional AS-PCR of the WT codon for detection of potential heterozygosity (or polyclonality). Controls used for AS-PCR: WT—subcloned Jurkat T cells, M—previously sequence-confirmed Jurkat mutant i.e., Y192F mutant DNA for Y192F AS-PCR and Y192E mutant for Y192E AS-PCR (for AS-PCR of Y192, M serves as a negative control and contains DNA from a Y192F mutant). -, - - - - PCR mix without DNA sample. (C and D) Western blots show the expression of indicated proteins in various LCK mutated clones (WT—subcloned Jurkat T cells, F—LCK Y192F, E—LCK Y192E, KO—knockout) (C) and TSAd mutated clones (D). (E) Example of the use of flow cytometry to assess clonality of an LCK mutated clone. The mutated clone and the original WT clone were respectively stained with low and high concentrations of CTV and mixed together for anti-LCK staining and analysed by flow cytometry. (F) Founder effect on CRISPR-mutants. Graph shows surface CD3 and intracellular LCK expression monitored by flow cytometry as in Figure 2C, of LCK Y192F and LCK KO mutated clones derived from two different founder Jurkat clones. Median and range of median fluorescence intensity of mutated clones, normalized to the corresponding founder WT clone, are shown.
sometimes contained additional mutations (indels or substitutions, i.e., sequence 1 and 2 in Figure 5B and sequence 1 in Figure 5C) in the vicinity of the targeted locus.

3.6 Differing knock-in efficiency in the LCK and the SH2D2A genes

A total of three different knock-in mutations in LCK at two different loci (the codons encoding Y192 and Y505 respectively) were performed in Jurkat T cells. The mutation frequency as estimated by AS-PCR after limiting dilution was between 5%-13% of the total number of clones screened (Figure 6A). Subsequent genotyping by sequencing revealed that nearly 80% of the AS-PCR positive clones were non-homozygous mutants (heterozygous or hemizygous) or knockout giving a false-positive signal (Figure 6B). Among the total number of clones isolated by limiting dilution for each independent experiment, on average 2% harboured the Y192E and Y192F and 1% harboured the Y505F homozygous knock-in mutation (Figure 6C).

CRISPR/Cas9-mediated cleavage of DNA is independent of whether the strand is sense or non-sense with respect to

---

**FIGURE 5** Screening knock-in clones with sequencing. Jurkat T cells were subjected to knock-in mutations for Y192E/F and Y505F in LCK. Amplicons of the region around the targeted loci from selected clones were subjected to Sanger sequencing. (A) Representative sequencing results showing various outcomes of the mutation. (B-C) Examples of various genome editing outcomes revealed by sequencing: (B) Sequences from cells targeted by LCK Y192 sgRNA (aberrant mutation outcomes are labelled 1-20) and (C) targeted by LCK Y505 sgRNA (aberrant mutation outcomes are labelled 1-7). Correctly introduced mutations are highlighted in yellow. Grey highlight represents unchanged WT sequence.
FIGURE 6 Efficiency of LCK and TSAd knock-ins. Jurkat T cells were subjected to knock-in mutations for Y192E/F and Y505F in LCK (A-C) or for R120K and Y290E/F in SH2D2A (D-F). Experiments, where symmetric or asymmetric ssODN were used, are indicated in open and filled symbols, respectively. (A and D) Graphs show percentage of AS-PCR positive clones for the indicated knock-in mutation experiments. Each point represents one independent experiment. \( n \) is displayed below the graph for each targeted mutation. The bars indicate mean values ± standard deviation. (B and E) Graph shows true homozygous mutants as revealed by sequencing for the indicated mutations. Total \( n \) of clones and \( n \) of the experiments for each knock-in mutation is indicated below the graph. (C and F) Graphs show percentage of true homozygous mutants as confirmed by sequencing for the indicated mutations. Each point represents one independent experiment. \( n \) is displayed below the graph. The bars indicate mean values ± standard deviation. (G and H) Epigenetic modifications of (G) LCK and (H) SH2D2A in primary T cells from peripheral blood. The exons targeted knock-in mutations in this study are indicated by the black arrows. The data sets are available at NIH Roadmap Epigenomics Mapping Consortium.\(^{30}\)
mRNA transcript for the respective proteins. After cleavage of DNA, the target strand remains bound to the DNA binding domain of Cas9, whereas the non-target DNA strand becomes exposed and available for pairing with the ssODN (Figure S1A-C). During the course of these studies, and as suggested by Richardson et al., we thus created asymmetric ssODN that were 103-127-nt long, with the shorter homology arm of the ssODN pairing with the non-target DNA strand 36-nt away from the cut site. We were, however, unable to detect an improved mutation frequency when using asymmetric ssODN for mutation of the same locus (Figure 6A,C).

In parallel efforts, we performed experiments to introduce four different knock-in mutations in three different loci (N52, R120 and Y290) in the SH2D2A gene. The knock-in mutation frequency after limiting dilution in Y290, as assessed by AS-PCR, was on the average 4% and 6% for the Y290E and Y290F mutations respectively (Figure 6D-F). Sanger sequencing of amplicons of the mutated regions revealed that on average 2% of the total number of clones screened were homozygous for the knock-in mutation. The homozygous knock-in frequencies for Y290E or Y290F were comparable to the frequencies of the knock-in mutations in the LCK gene (Figure 6B). In contrast, less than 1% of total clones were positive for R120K as determined by AS-PCR (Figure 6D). Of the two clones that were positive for the knock-in allele by AS-PCR, genotyping by Sanger sequencing did not reveal any homozygous knock-in mutations (Figure 6E,F). For mutation of N52 to S52, in two independent experiments we were unable to detect any successful knock-in mutations in bulk transfected Jurkat T cells prior to limiting dilution, neither did we observe any AS-PCR positive clones after limiting dilution (data not shown).

3.7 Epigenetic gene profiles correspond to knock-in efficiency

Our data show that the CRISPR/Cas9 knock-in mutation efficiency varied considerably both between genes and also between different loci within the same gene. Whereas LCK and TSAd are both expressed in T cells, LCK is constitutively expressed while TSAd expression is upregulated following T cell activation. We therefore speculated that the expression pattern of the genes could have implications for gene editing. CRISPR/Cas9 target cleavage may be inhibited by nucleosomes or other epigenetic mechanisms.

We hypothesized that SH2D2A is inaccessible to Cas9 in resting cells as a consequence of epigenetic modifications or local chromatin structure. The epigenetic profile of primary T cells from peripheral blood supports this notion (Figure 6G), as the four exons at the 5' end of SH2D2A, encoding S52 and R120, are associated with multiple histone markers. In contrast, the 3' exons of SH2D2A, including exon 7 which contains the Y290 codon (Figure 6G) has a different epigenetic profile, resembling that of LCK exon 6 and exon 13 encoding Y192 and Y505 respectively (Figure 6H).

In an attempt to increase the accessibility of Cas9 to the TSAd locus, we activated Jurkat T cells by anti-CD3 or PMA/ionomycin (treatments which induce expression of TSAd in Jurkat T cells) before or after transfection for gene editing of TSDa R120. Only one TSAd R120K heterozygous clone was obtained following activation of the Jurkat cells with anti-CD3 antibody prior to CRISPR/Cas9 transfection. All TSAd R120 mutation approaches described here are included in Figure 6D-F. Although the generation of the TSAd R120K heterozygous mutant may be due to pre-activation, it is also possible that this was a result of a Cas9-independent HDR event.

4 DISCUSSION

In this study, we report our experience with gene editing of Jurkat T cells to generate mutated cell lines which can be used to study domain-specific functions of signalling molecules in T cell activation. Further, we report locus-dependent knock-in mutation efficiencies, also within the same gene, which may be due to local variation in chromatin structure. Our workflow for generation of knock-in mutants using the CRISPR/Cas9 system can easily be applied by laboratories equipped to use basic molecular biology techniques.

To successfully introduce a desired knock-in mutation using CRISPR/Cas9, there are two events which must occur concurrently in a precise manner. Firstly, the Cas9 machinery must be targeted to the correct location, and secondly, the Cas9 endonuclease must facilitate DNA cleavage. This involves accurate directing of Cas9 by the sgRNA to the target site and efficient binding to the PAM site respectively.

Unsuccessful CRISPR/Cas9 experiments may be caused by PAM site inaccessibility. Several groups have addressed ways to improve Cas9 target efficiency, for instance by using algorithms to predict efficient sgRNA target sites, according to sequence features. In the case of increasing knock-in efficiency, selecting efficient PAM sites that are < 10-nt away from the site of the mutation, was proposed.

The limited availability of PAM sites near the mutation site represents an important challenge for successful generation of knock-in mutations. The frequency of ‘GG’ in the human genome has been reported to be 5.21%. This means that ‘GG’ should occur on average every 42 bases. Accounting also for ‘CC’ (representing the PAM site on the complementary strand) will give on average one PAM site within a 10 base distance from the mutagenic site of interest. Although PAM site recognized by Cas9 from S pyogenes is relatively simple and abundant, it might be advantageous to use CRISPR systems from other bacteria, which recognize
other PAM sites.\textsuperscript{35} It could potentially allow cutting of DNA closer to the target site, and increase the efficiency of HDR.

Since the first reports of gene disruption in Jurkat T cells using CRISPR/Cas9,\textsuperscript{36,37} there has been an ever-increasing number of reports where CRISPR/Cas9 technology has been used to disrupt selected genes in Jurkat T cells (80 publications in PubMed, September 2019). However, to our knowledge, the efficiency of gene editing by HDR in Jurkat T cells using CRISPR/Cas9 technology has not previously been reported. The relatively low homozygous knock-in frequency compared to knockout frequencies observed in our study of Jurkat T cells, is consistent with previously reported data from human primary CD4$^+$ T cells and CD34$^+$ HSPCs.\textsuperscript{38} Although the frequencies were low, we repeatedly introduced knock-in mutations across independent experiments with the same experimental set-up (Figure 6A-C). We have successfully used the resulting mutants to study LCK function in T cell signalling (Borowicz et al., submitted).

Generation of knock-in mutants is time-consuming, and one single experiment takes a minimum of three weeks to complete. It is therefore advantageous to check the performance of the sgRNA already prior to the limiting dilution step, while the cells are still in bulk. However, this strategy also has its limitations. If detection primers are complementary to the ssODN repair template, initial bulk DNA analysis may give false-positive results, as the ssODN may still be present in the medium or the transfected cells. This should however not be a problem during the clonal screening, as the ssODN would be sufficiently diluted or degraded by then. Hence, if possible, the detection primers should be non-complementary to the ssODN.

An alternative to testing sgRNA and/or the knock-in mutation efficiency with AS-PCR would be to sequence amplicons from mutated bulk cells and analyse by software such as ICE\textsuperscript{39} or TIDE.\textsuperscript{40} However, in our opinion, developing an AS-PCR set-up for screening of a large number of clones with possible knock-in mutations is generally easy to do and more affordable than Sanger sequencing.

If the targeted site is located in the C-terminus of the protein (as in the case of LCK Y505), the potential mutations may not generate protein knockout, but rather truncated versions of the protein. Additionally, it is important to be aware of the existence of alternative spliced variants of the targeted protein. If the targeted mutation is located in an exon that can be omitted by alternative splicing, the mutation will not deplete the cell of all splicing variants of the targeted gene (as exemplified by TSAd Y290 in exon 7 of SH2D2A).\textsuperscript{41}

CRISPR/Cas9-mediated genome editing involves a certain risk of generating off-target effects in the treated cells due to unspecific guidance of Cas9.\textsuperscript{42} However, the heterogeneity of the resulting mutated clones may also be a consequence of the inherent heterogeneity of the starting cell lines, as our analysis of the Jurkat T cell subclones clearly demonstrated (Figure 2A,C). To control pre-existing heterogeneity, we thus strongly recommend generating one or several founder cell lines. It is particularly important to ensure sufficient expression of key molecules crucial in pathways contributing to phenotypes in the knock-in cell line. In our case, establishing founder Jurkat T cell lines was necessary to ensure that the phenotypes of the later knock-in mutations were not biased by the variation in the expression of surface CD3, LCK or TSAad between the mutated clones. Comparing mutants derived from at least two independent founder cells (ie data presented in Figure 4F), can reveal hidden heterogeneity between the founders which may influence the mutant phenotype. It will help to determine whether an observed phenotype is a result of the mutation or if it is related to the founder cell line background.

An additional source of unspecific effects of CRISPR/Cas9-mediated gene editing is the stochastic nature of DNA repair mechanisms, which can create substantial local gene rearrangements\textsuperscript{43} as we also observed (Figure 5B,C). A striking observation in our experiments was repeated similar repair patterns observed in independent mutation experiments. This fits with the notion that particular outcomes of NHEJ DNA repair to some extent can be anticipated. The software InDelphi,\textsuperscript{44} predicts the likelihood of given NHEJ results based on microhomology analysis. We have compared our results with InDelphi predictions in Figure S5. Consistent with the InDelphi predictions, we observed that the sequences predicted to occur with the highest frequencies were observed in our sequencing results.

Our workflow proposes methods for mutated clones' validation, which should ensure desired homozygous knock-in mutations and omit clones with unspecific mutations in the targeted protein. In some of our knock-in mutants, we observed protein products of various lengths, which were not present in the WT samples. Even if the expressed knock-in mutated protein seems to be of correct length by Western blotting, DSB increases the chance of introducing random mutations into the gene of interest. Therefore, it should be carefully monitored by sequencing. We observed within the 250 bp sequenced, additional mutations occurring at some distance from the targeted codon. Thus, it is possible that we may have overlooked some additional mutations outside of the sequenced region.

A number of sgRNA-designing online software\textsuperscript{20} also allow prediction of potential off-target effects outside of the target gene. Sequencing of these potential off-target sites would further ensure that the mutant's phenotype is the result of the specific mutation and not the result of other genes' mutations.
A striking observation from our work was the variable success rate for CRISPR/Cas9 targeting at different loci in the same gene, which might be explained by variable chromatin accessibility. This could be due to an error in the generation of the sgRNA or Cas9 from our px330 plasmid. However, for the R120 site in SH2D2A, we tested three different PAM sites and sgRNA target sites to improve efficiency, while still remaining in proximity to the desired site of mutation. Still, we were unable to detect any successful knock-in mutations. We thus hoped that activation of the Jurkat T cells prior to the mutation experiments, would change local chromatin structure and hence increase the accessibility of Cas9 to the target mutation site. However, despite different activation methods, we were unable to successfully increase the frequency of knock-in mutations at the R120 site and only obtained one heterozygous clone for R120K.

Although activation of Jurkat T cells through TCR prior to gene editing did not improve the mutation efficiency of TSAd R120, it does not exclude the possibility that local chromatin structure is a determining factor. Besides Cas9 accessibility, local chromatin domains might determine which DNA repair pathway is taken to repair the DSB, that is through NHEJ or HDR. Furthermore, it is known that NHEJ occurs more frequently and at a faster rate compared to HDR, which may result in HDR machinery being outcompeted by NHEJ. If this is the case, we would expect to see a high frequency of indels at the R120 site, which to a certain degree can be picked up by our screening methods. As we did not observe any abnormalities, it suggests that R120 was not successfully targeted by the sgRNA and Cas9. Several groups have addressed ways to promote HDR pathways, such as calibrating the cells to be in the G2 phase of the cell cycle, where HDR is suggested to be more active than NHEJ repair mechanisms. Alternatively, inhibiting NHEJ pathways might be a possible strategy to enhance HDR-mediated knock-in efficiency.

In conclusion, we have successfully used CRISPR/Cas9 technology to target single domains in the LCK and SH2D2A genes using a set-up which is simple and easy to establish for labs with basic skills in molecular biology. Our work reveals variability in mutation frequency, also within the same gene. Inability to target a given locus may be associated with the chromatin structure in the vicinity of the locus affecting sgRNA abilities to guide Cas9 to the desired genomic location.

ACKNOWLEDGMENT
This work was supported by the University of Oslo, the Norwegian Cancer Society (grant 107561 - PR-2007-0193), Novo Nordisk Fonden, Unifor and Anders Jahres fond til vitenskapsfremme.

CONFLICT OF INTEREST
The authors declare no competing interests.

AUTHOR CONTRIBUTIONS
AS, PB and HC conceptualized the study. AS, PB, HC, DM and SG designed experiments and analysed results. PB performed Western Blotting and flow cytometry experiments. PB, HC DM and SG generated mutated and founder cell lines. PB deconvoluted mutated sequences. HK performed molecular cloning of all necessary constructs. AS, PB and HC wrote the manuscript with the input from all co-authors. AS oversaw the project.

ORCID
Pawel Borowicz https://orcid.org/0000-0002-5604-5844
Hanna Chan https://orcid.org/0000-0002-7822-8262
Anne Sparkland https://orcid.org/0000-0003-4421-0766

REFERENCES
1. Urnov FD. Genome editing BC (before CRISPR): last lessons from the “old testament”. CRISPR J. 2018;1:34-46.
2. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315:1709-1712.
3. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281-2308.
4. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013;339:819-823.
5. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157(6):1262-1278.
6. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol. 2014;32:347-355.
7. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna J, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337:816-821.
8. Mao Z, Bozella M, Seluanaov A, Gorbonova V. Comparison of nonhomologous end joining and homologous recombination in human cells. DNA Repair. 2008;7:1765-1771.
9. McVey M, Lee SE. MMEJ repair of double-strand breaks (director’s cut): deleted sequences and alternative endings. Trends Genet. 2008;24:529-538.
10. Shalem O, Sanjana NE, Hartenian E, et al. Genome-scale CRISPR-Cas9 knockout screening in human cells. Science. 2014;343:84-87.
11. Sternberg SH, Doudna JA. Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol Cell. 2015;58:568-574.
12. Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol. 2016;34:339.
13. Granum S, Sundvold-Gjerstad V, Gopalakrishnan RP, et al. The kinase Itk and the adaptor TSAd change the specificity of the kinase Lck in T cells by promoting the phosphorylation of Tyr192. Sci Signal. 2014;7:ra118.
14. Granum S, Andersen TCB, Sortlie M, et al. Modulation of Lck function through multisite docking to T cell-specific adapter protein. J Biol Chem. 2008;283:21909-21919.
15. Kaplun L, Fridman AL, Chen W, et al. Variants in the signaling protein TSAd are associated with susceptibility to ovarian
cancer in BRCA1/2 negative high risk families. Biomark Insights. 2012;7:151-157.
16. Sundvold V, Torgersen KM, Post NH, et al. T cell-specific adapter protein inhibits T cell activation by modulating Lck activity. J. Immunol. 2000;165:2927-2931.
17. Spurkland A, Brinchmann JE, Markussen G, et al. Molecular cloning of a T cell-specific adapter protein (TSAd) containing an Src homology (SH) 2 domain and putative SH3 and phosphotyrosine binding sites. J Biol Chem. 1998;273:4539-4546.
18. Sundvold-Gjerstad V, Granum S, Mustelin T, et al. The C terminus of T cell-specific adapter protein (TSAd) is necessary for TSAd-mediated inhibition of Lck activity. Eur J Immunol. 2005;35:1612-1620.
19. Clippe NA, Crabtree GR. Identification of calcineurin as a key signalling enzyme in T-lymphocyte activation. Nature. 1992;357:695-697.
20. Haeussler M, Schönig K, Eckert H, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 2016;17:148.
21. Koshy L, Anju AL, Harikrishnan S, et al. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res. 1990;18:999-1005.
22. Snow K, Judd W. Heterogeneity of a human T-lymphoblastoid cell line. Exp Cell Res. 1987;171:389-403.
23. Ugozzoli L, Wallace RB. Allele-specific polymerase chain reaction. Methods. 1991;2:42-48.
24. Gaudet M, Fara AG, Bertognolo I, Sabatti M. Allele-specific PCR in SNP genotyping. Methods Mol Biol. 2009;578:415-424.
25. Kwok S, Kellogg DE, McKinney N, et al. Effects of primer-template mismatches on the polymerase chain reaction: human immunodeficiency virus type 1 model studies. Nucleic Acids Res. 1990;18:999-1005.
26. Huang M-M, Arnheim N, Goodman MF. Extension of base mispairs by Taq DNA polymerase: implications for single nucleotide discrimination in PCR. Nucleic Acids Res. 1992;20:4567-4573.
27. Lajin B, Alachkar A, Alhaj Sakur A. Betaine significantly improves structure function of CRISPR-Cas9. Nature. 2016;533.
28. Scherer S. Short Guide to the Human Genome. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2008.
29. Wang Y, Liu KI, Sutrisno N-AB, et al. Systematic evaluation of CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). J Biol Chem. 1992;273:4539-4546.
30. Bernstein BE, Stamatoyannopoulos JA, Costello JF, et al. The NIH Roadmap Epigenomics Mapping Consortium. Nat Biotechnol. 2010;28:1045-1048.
31. Uusi-Mäkelä MIE, Barker HR, Bäuerlein CA, Häkkinen T, Nykter M, Hiilimetsä M. Chromatin accessibility is associated with CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). PLoS One. 2018;13:e0196238.
32. Doench JG, Fusi N, Sullender M, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol. 2016;34:184-191.
33. Paquet D, Kwart D, Chen A, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR-Cas9. Nature. 2016;533.
34. Scherer S. Short Guide to the Human Genome. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2008.
35. Wang Y, Liu KI, Sutrisno N-AB, et al. Systematic evaluation of CRISPR-Cas9 efficiency in the zebrafish (Danio rerio). J Biol Chem. 1992;273:4539-4546.
36. Navarro JM, Touzart A, Pradel LC, et al. Site- and allele-specific polycomb dysregulation in T-cell leukaemia. Nat Commun. 2015;6:6094.
37. Neggers JE, Vercruysse T, Jacquemyn M, et al. Identifying drug-target selectivity of small-molecule CRM1/XPO1 inhibitors by CRISPR/Cas9 genome editing. Chem Biol. 2015;22:107-116.
38. Mandal PK, Ferreira LMR, Collins R, et al. Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. Cell Stem Cell. 2014;15:643-652.
39. Hsiau T, Maures T, Waite K, et al. Inference of CRISPR edits from Sanger trace data. BioRxiv. 251082 (2018).
40. Brinkman EK, van Steensel B. Rapid Quantitative Evaluation of CRISPR Genome Editing by TIDE and TIDER. New York, NY: CRISPR Gene Editing. Humana Press; 2019:29-44.
41. Granum S, Sundvold-Gjerstad V, Dai KZ, et al. Structure function analysis of SH2D2A isoforms expressed in T cells reveals a crucial role for the proline rich region encoded by SH2D2A exon 7. BMC Immunol. 2006;7:15.
42. Fu Y, Foden JA, Khayer C, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol. 2013;31:822-826.
43. Kosicki M, Tomberg K, Bradley A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. Nat Biotechnol. 2018;36:765.
44. Shen MW, Arbab M, Hsu JY, et al. Predictable and precise template-free CRISPR editing of pathogenic variants. Nature. 2018;563:646-651.
45. Gu K, Rehman S, Liu M, et al. Methodologies for Improving HDR Efficiency. Front Genet. 2019;9:1-9.
46. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol. 2015;33:538-542.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.