Efficient Human Cell Coexpression System and Its Application to the Production of Multiple Coronavirus Antigens

Chonsaeng Kim, You Kyeong Jeong, Jihyeon Yu, Hye Jin Shin, Keun Bon Ku, Hyung Jin Cha, Jun Hee Han, Sung-Ah Hong, Bum-Tae Kim, Seong-Jun Kim,* Jae-Sung Woo,* and Sangsu Bae*

Coproduction of multiple proteins at high levels in a single human cell line would be extremely useful for basic research and medical applications. Here, a novel strategy for the stable expression of multiple proteins by integrating the genes into defined transcriptional hotspots in the human genome is presented. As a proof-of-concept, it is shown that EYFP is expressed at similar levels from hotspots and that the EYFP expression increases proportionally with the copy number. It is confirmed that three different fluorescent proteins, encoded by genes integrated at different loci, can be coexpressed at high levels. Further, a stable cell line is generated, producing antigens from different human coronaviruses: MERS-CoV and HCoV-OC43. Antibodies raised against these antigens, which contain human N-glycosylation, show neutralizing activities against both viruses, suggesting that the coexpression system provides a quick and predictable way to produce multiple coronavirus antigens, such as the recent 2019 novel human coronavirus.

1. Introduction

Purified recombinant proteins are extremely useful for basic research in biology and chemistry as well as in medical applications involving therapeutic proteins. To obtain proteins with high purity and yield, their overexpression is essential. Therefore, various heterologous expression systems including *Escherichia coli*, yeast, insect cell, and mammalian cell systems have been developed for high-level protein expression. Among them, mammalian expression systems have received special attention because they could efficiently produce many

recombinant proteins that other systems could not.[1] For examples, Chinese hamster ovary (CHO) cells are widely used for the production of therapeutic proteins such as monoclonal antibodies and hormones,[2] and human embryonic kidney 293 (HEK-293) cells are increasingly used in biochemistry and structural biology fields for the production of human proteins that require the proper cellular factors and environments for expression, folding, modification, or secretion. In particular, proteins undergo human-like post-translational modifications (PTMs), which are frequently critical for proper protein structure, activity, stability, optimal pharmacokinetics, and immunogenicity.[3,4] However, the use of the human cell system is often limited due to inconsistent and unpredictable productivity of transient and stable transfection methods, respectively. Therefore, a new method to systemically increase the expression level in human cells would be beneficial.

A reliable coexpression method for the simultaneous expression of multiple proteins greatly increases the usability of an expression system. It is crucial to produce unstable proteins that need to be expressed with a binding partner as well as cytotoxic proteins that have to be expressed with an inhibitory protein with a higher expression level at the same cellular location. In addition, it is beneficial, in terms of time, cost, and labor, for the coexpression of multiple viral antigens or large

Dr. C. Kim, Dr. H. J. Shin, K. B. Ku, Dr. B.-T. Kim, Dr. S.-J. Kim
Center for Convergent Research of Emerging Virus Infection
Korea Research Institute of Chemical Technology
Daejeon 34114, South Korea
E-mail: sekim@krict.re.kr

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/adbi.202000154.

Y. K. Jeong, Dr. J. Yu, J. H. Han, S.-A. Hong, Prof. S. Bae
Department of Chemistry
Hanyang University
Seoul 04763, South Korea
E-mail: sangsubae@hanyang.ac.kr

Y. K. Jeong, Dr. J. Yu, J. H. Han, S.-A. Hong, Prof. S. Bae
Research Institute for Convergence of Basic Sciences
Hanyang University
Seoul 04763, South Korea

Dr. H. J. Cha, Prof. J.-S. Woo
Department of Life Sciences
Korea University
Seoul 02841, South Korea
E-mail: jaesungwoo@korea.ac.kr

protein complexes, even when the antigens or subunits can be separately expressed and purified. To address this issue, various coexpression methods have been developed in several species. For examples, in *E. coli* expression systems, multiple expression cassettes independently controlled by separate promoters can be inserted into a single plasmid using conventional cloning methods,[5] or several vectors conferring resistance to different antibiotics can be cotransformed into a single *E. coli* cell.[6] Likewise, in the Multibac system for coexpression in insect cells, a baculovirus bacmid containing six expression cassettes can be generated by Cre-Lox recombination.[7] However, there is still no reliable coexpression method for mammalian cells including human cells, although mammalian cell systems are used as widely as other systems. Multiple plasmids can be cotransfected, but this approach typically results in extremely various expression levels of the target proteins in individual cells, because the relative plasmid transfection efficiencies vary greatly from cell to cell. Furthermore, generating a single coexpression plasmid containing multiple expression cassettes is also difficult, because repeated transcription-enhancing elements and polyalanine-encoding sequences greatly decrease plasmid stability during cloning and amplification in *E. coli*. Therefore, a new method that ensures high-level expression of multiple genes in human cells is highly desirable.

A predictable and stable human cell expression system is required for producing the most effective recombinant proteins for human vaccines, which are one of the most practical precautions against infection by viruses, including coronavirus. To prophylactically stimulate the human immune system, effective antigens that do not themselves cause disease must be delivered to the human body. In contrast to live attenuated or inactivated vaccines that involve the whole pathogen, subunit vaccines based on individual recombinant viral proteins reduce the possibility of adverse reactions. However, recombinant subunit vaccines have potential challenges, such as a weak immunogenicity and the need for immune-boosting adjuvants and multiple doses.[8] Previous works have indicated that human-like PTMs including glycosylation are crucial for the immunogenicity of recombinant antigens in humans,[4,9–13] emphasizing the need for reliable and effective systems for expressing recombinant antigens in human cells.

Here, we developed a general method for the stable expression of multiple recombinant proteins in human cells by integrating genes of interest into several transcriptional hotspots in the human genome. To integrate target genes into the hotspots specifically, we harnessed CRISPR (clustered regularly interspaced short palindromic repeat)-Cas (CRISPR associated)-based knock-in technologies.[14–20] We verified each transcriptional hotspot using fluorescent protein coding genes as a proof-of-concept and confirmed that three different fluorescent protein coding genes integrated at different loci can be coexpressed at high levels. We further established clonal cell lines in which two different antigens were continuously expressed at similar levels and ultimately showed that the two purified antigens effectively induced potent neutralizing activities against two viruses, suggesting a novel and cost-effective strategy for vaccine production.

2. Results

2.1. Precise Knock-in Strategies at Transcriptional Hotspots in the Human Genome

We first examined transcriptional hotspots, identified in previous studies, in which integrated genes are abundantly expressed because of open chromatin structures and at which the integration does not affect neighboring gene expression.[21] We carefully selected seven hotspots in different chromosomes, including the first intron of the *PPP1R12C* gene, which is referred to as Adeno-associated virus integration site 1 (AAVS1)[21–24] (**Figure 1**a). We next investigated whether genes integrated at these hotspots showed high-level expression. To this end, we used enhanced yellow fluorescent protein (EYFP) gene as a surrogate, so that we could assess the expression level by measuring the fluorescence Intensity of the cell.

To precisely integrate the EYFP gene into each hotspot region, we sought to establish two different knock-in strategies: a homology directed repair (HDR)-mediated knock-in method and a homology-independent targeted insertion (HITI) method that is based on the nonhomologous end joining (NHEJ) DNA repair pathway.[25–27] We initially constructed a donor DNA template that contains the EYFP gene with expression driven by an inducible cytomegalovirus (CMV) promoter and the hygromycin resistance (HygR) gene for effective selection of knock-in cells (Figure 1b). Then, for HDR-mediated knock-in, the donor plasmid was designed to contain two homology arm regions, flanking the HygR and EYFP genes, which are identical to 800 bp regions upstream and downstream of a cleavage site that will be induced at a target site of interest. In contrast, one or both of the flanking region(s) in the donor plasmid for HITI contained a 23-bp sequence identical to the cleavage sequence (Figure S1, Supporting Information).

To generate DNA cleavage at specific sites of interest, we designed at least five single-guide RNAs (sgRNAs) targeting sites in each hotspot using Cas-Designer and Cas-OFFinder software[28,29] (Figure S2, Supporting Information). We initially focused on sites recognized by SpCas9 (Cas9 derived from *Streptococcus pyogenes*), which recognizes 5'-NGG-3' protospacer-adjacent motif (PAM) sequences, for all hotspots. In the case of the SPINK2 hotspot region, we found that a long noncoding RNA (lncRNA) was encoded downstream of the SPINK2 gene, so we designed an additional five sgRNAs that avoided sites in the lncRNA gene. It is notable that other programmable nucleases, including zinc-finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and other CRISPR endonucleases such as Cas12a (also called Cpf1), can also be used for specific DNA cleavage in hotspot regions.[30–34] Hence, we alternatively designed sgRNAs for LbCas12a (Cas12a derived from *Lachnospiraceae* bacterium ND2006), which recognizes 5'-TTTV-3' PAM sequences (Figure S3, Supporting Information).[35]

2.2. Verification of Transcriptional Hotspots in the Human Genome

To test the gene editing activities of all of the sgRNAs, we transfected plasmids encoding SpCas9 or LbCas12a and, individually,
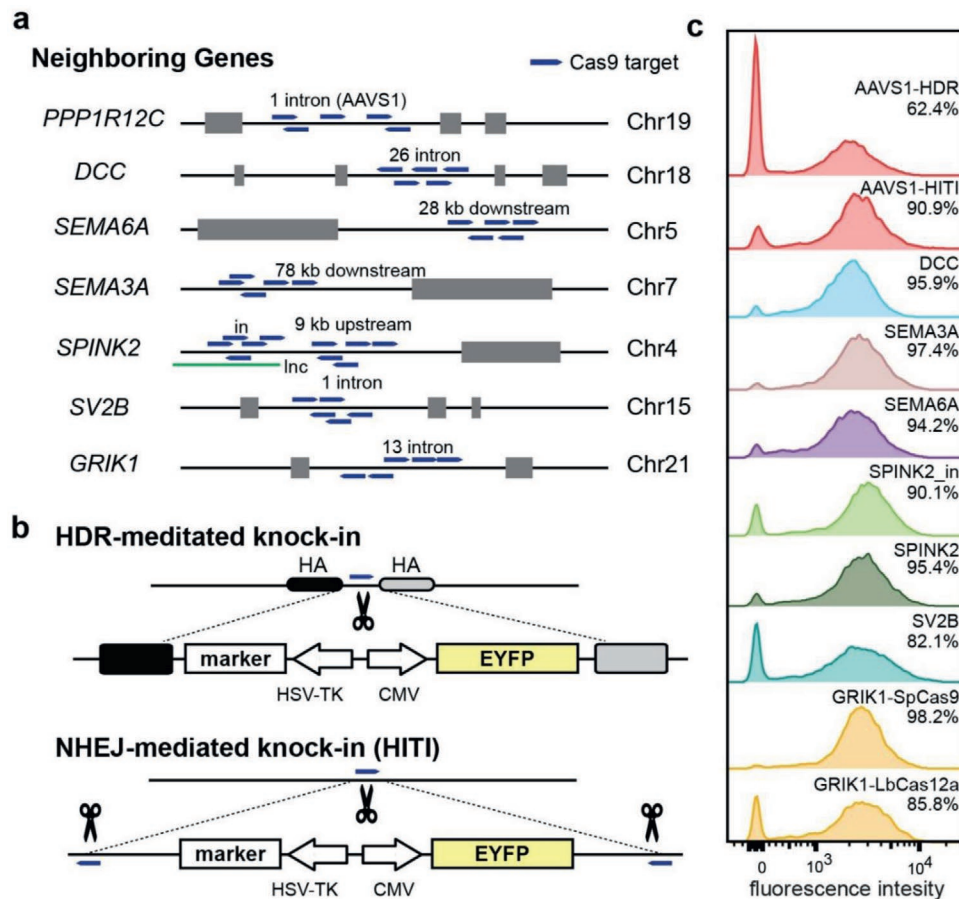
Figure 1. Determination of human transcriptional hotspots. a) Positions of human transcriptional hotspots based on neighboring gene information. The grey boxes indicate exons in the neighboring genes and the blue boxes with an arrowhead on one end indicate selected SpCas9 targets. The green bar represents the position of a lncRNA. b) Schematic diagram of two integration methods using programmable endonucleases. In each case, the upper bar indicates the targeted sites within the genome; the linked bar underneath indicates the DNA template. The DNA templates contain two promoters, which control the expression of a gene encoding a fluorescent protein (EYFP) and an antibiotic resistance gene (marker). HA, homology arm. c) Histogram plots of flow cytometry analysis. The fluorescence intensity of each population of cells was characterized. The height of each plot indicates the cell count. Without AAVS1, the neighboring genes of (a) are called transcriptional hotspots instead.

each of the sgRNAs into human T-REx-293 cells.[36] We then used targeted deep sequencing to measure the frequency of small insertions and deletions (indels) at the target site, a measure of sgRNA activity.[37] We chose the sgRNAs showing the highest activities for each of the hotspots from the high-throughput sequencing data (Figures S4–S6, Supporting Information). Next, we tested HDR- and HITI-mediated knock-in of the EYFP gene at the AAVS1 hotspot by transfecting the EYFP-encoding HDR or HITI donor plasmid, together with plasmids encoding SpCas9 and the AAVS1-targeted sgRNA, into the cells. Transfected cells were selected under hygromycin treatment and flow cytometry was used to detect EYFP-positive cells. The results showed that the fluorescence intensities were comparable for each of the two knock-in methods (Figure 1c and Figure S7, Supporting Information), indicating that the knock-in method did not affect the expression level of the integrated gene.

After successful integration of the EYFP gene at the AAVS1 site, we independently integrated the EYFP gene into all of the other hotspot sites using the HITI method. As described above,

after hygromycin selection, cells were analyzed by flow cytometry to compare EYFP expression levels from each of the hotspots with that from the AAVS1 site (Figure 1c, Figures S7 and S8, Supporting Information). We found that all cell populations exhibited equivalently high expression levels, indicating that all selected sites can be used as transcriptional hotspots. Furthermore, we integrated the EYFP gene into the GRIK1 region using LbCas12a nuclease in addition to SpCas9 and found that the fluorescence intensities were almost the same (Figure 1c, Figures S7 and S8, Supporting Information), suggesting that other programmable nucleases, including ZFNs and TALENs, could be used as the knock-in tool.

2.3. Multiplexed Target Gene Expression from Human Transcriptional Hotspots

We next investigated whether multiple genes can be simultaneously expressed from distinct hotspots. To insert genes encoding three different fluorescent proteins into different
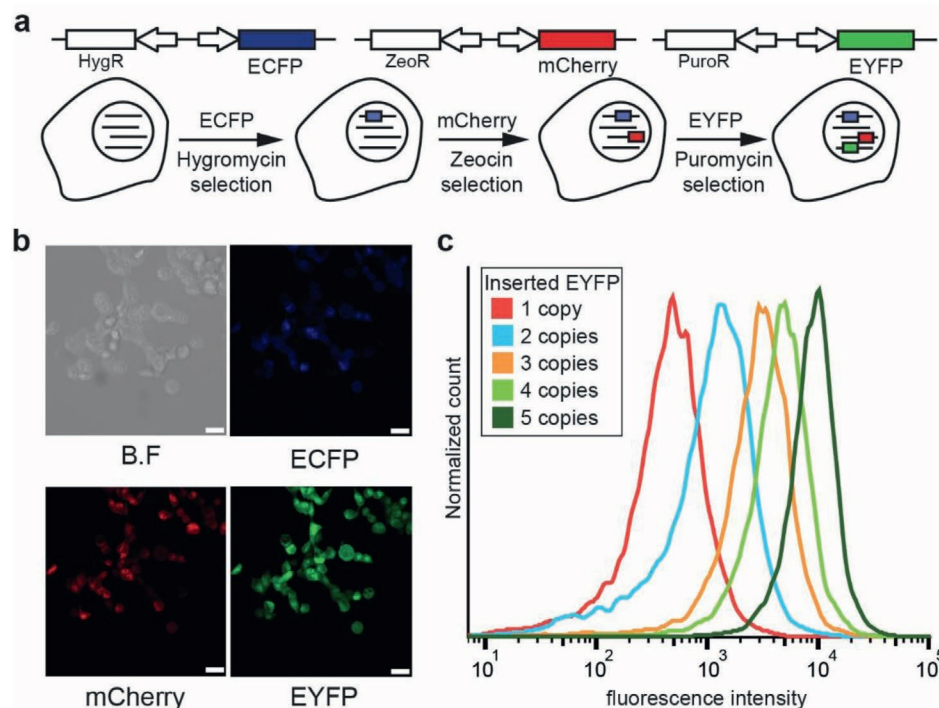
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
BIOLOGY**

www.advanced-bio.com

Figure 2. Ability to multiplex target gene expression at human transcriptional hotspots. a) Schematic diagram showing the multistep generation of the desired cell populations. Three kinds of DNA templates, each containing a gene encoding a different fluorescent protein and a different antibiotic resistance gene, were constructed. In turn, the templates were integrated into distinct transcriptional hotspots, after which the cells were treated with the appropriate antibiotic (hygromycin for integration at the AAVS1 site, zeocin for integration at the SEMA6A site, and puromycin for integration at the DCC site). b) Fluorescence confocal microscope images of the final cell population generated in (a). White scale bars, 10 μm. c) EYFP encoding genes were sequentially integrated in transcriptional hotspots (AAVS1, SEMA6A, GRIK1, and SEMA3A). During sequential integration steps, cell lines containing 1–5 copies of EYFP encoding genes were separated by clonal selection. The fluorescence intensity of the clonal cell lines was characterized in merged histogram plot. The height of each plot was normalized with maximum value of the peak.

hotspots, we constructed three donor plasmids, as follows: the first contains the enhanced cyan fluorescent protein (ECFP) gene and the HygR gene for hygromycin resistance, the second contains the mCherry (red fluorescent protein) gene and the Sh ble (*Streptoalloteichus hindustanus* bleomycin) gene for Zeocin resistance, and the third contains the EYFP gene and the pac (puromycin N-acetyl-transferase) gene for puromycin resistance; they are designed for integration at the AAVS1, SEMA6A, and DCC hotspots, respectively (Figure S9, Supporting Information). Then, we sequentially transfected the donor plasmids, along with plasmids encoding SpCas9 and the appropriate sgRNA, into human T-REx cells (**Figure 2**a), to successively obtain ECFP knock-in cells, ECFP and mCherry knock-in cells, and ECFP, mCherry, and EYFP knock-in cells. We confirmed the presence of the ECFP, mCherry, and EYFP proteins in the final cell population by detecting fluorescence at the expected wavelengths with confocal fluorescence microscopy (Figure 2b, Figures S10 and S11, Supporting Information), indicating that the three different proteins can be expressed in one stable cell line.

To investigate whether the integration of multiple copies of a single target gene would proportionately increase the expression level, we further prepared various EYFP-encoding donor plasmids that contain different antibiotic resistance genes. Then, we sequentially introduced each plasmid at different hotspots. Ultimately, we constructed a number of clonal cell

lines that contain from one to five copies of the EYFP gene (Figures S12 and S13, Supporting Information). For each cell line, we assessed the EYFP expression level by measuring fluorescence intensities by flow cytometry (Figure 2c and Figure S14, Supporting Information). Interestingly, the fluorescence intensities increased according to the number of copies of the EYFP gene, indicating that hotspot-based gene expression can be multiplexed.

2.4. Expression of an MERS Immunogen in Human Cells

We next applied the human coexpression system to generate a recombinant subunit vaccine. As an antigen, we chose the receptor-binding domain (RBD) of the Middle East respiratory syndrome coronavirus (MERS-CoV) spike protein (**Figure 3**a). The spike-RBDs, which cover the outer surface of coronaviruses, are responsible for binding to the MERS-CoV receptor, dipeptidyl peptidase 4 (DPP4, also known as CD26) (Figure 3b), and subsequent fusion with a host cell membrane. A previous study showed that MERS-RBD fused with human Fc, produced by transient expression in HEK-293 cells, induced potent neutralizing antibodies in a mouse model.[38] We introduced the sequence encoding this MERS-RBD-human Fc fusion antigen into the donor expression plasmid for HDR-mediated knock-in. The signal peptide-encoding sequences of the MERS-CoV spike
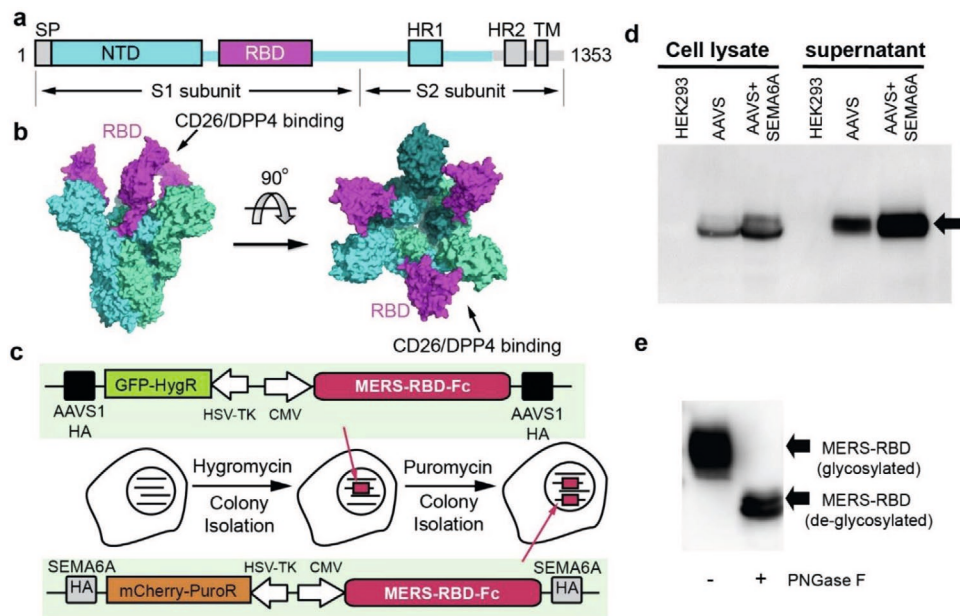
**ADVANCED
SCIENCE NEWS**
www.advancedsciencenews.com

**ADVANCED
BIOLOGY**
www.advanced-bio.com

Figure 3. Expression of MERS antigen in human cells. a) Schematic representation of the MERS-CoV spike protein. SP, signal peptide; NTD, N-terminal domain; RBD, receptor-binding domain; HR1, heptad repeat 1; HR2, heptad repeat 2; TM, transmembrane domain. b) Structure of the trimeric MERS-CoV spike protein in the pre-fusion conformation. The atomic coordinates were obtained from the Protein Data Bank (PDB code: 5×59). Surface presentations were generated using PyMol program. The three protomers, with the exception of the RBD domain (magenta), are colored cyan, deep teal, and green-cyan. c) Schematic diagram of vectors expressing the MERS antigen and generation of stable cell lines expressing this antigen using CRISPR knock-in technology. d) Cell lines in which the desired construct(s) were inserted in the AAVS hotspot and in both the AAVS and SEMA6A loci were cultivated. Cells (cell lysate) and culture supernatants (supernatant) were analyzed by Western blot using an antihuman Fc antibody. e) Western blot analysis of the MERS-RBD (detected with an anti-His antibody) before and after treatment with PNGase F, which cleaves N-linked glycans. Arrows indicate the glycosylated and deglycosylated forms of the MERS antigen.

protein were fused to the antigen-encoding sequences, because these peptides induce the PTMs and secretion of the protein into the culture supernatant, facilitating antigen purification without host cell lysis. Then, we sequentially introduced the MERS antigen-expressing donor plasmid into the AAVS1 and SEMA6A loci (Figure 3c and Figure S15, Supporting Information). After antibiotic selection and colony isolation, cell lines strongly expressing the antigen were selected. We checked the MERS antigen level both in cell lysates and in cell supernatants using an antihuman Fc antibody (Figure 3d), and observed that an increase in the number of hotspot-integrated genes from one to two resulted in a proportionate increase in antigen expression (with the antigen yield increasing from 7.3 to 14.8 mg L^{-1}). We also found that treatment with PNGase F, which removes N-linked oligosaccharides, resulted in a large shift of the protein band on an SDS-PAGE gel (Figure 3e). This result suggests that the purified MERS antigens are heavily N-glycosylated.

2.5. Simultaneous Expression of MERS-CoV and Human Coronavirus OC43 Immunogens

Multiple antigens from one or more pathogens can be more effective and/or cost-effective than a single antigen.[39,40] Hence, we next aimed to make a stable cell line that produces multiple subunits, which would in turn be used as a vaccine that simultaneously targets two or more virus species. To this end, we selected two antigens from different species, MERS-CoV and

human coronavirus OC43 (HCoV-OC43); both belong to the same *Coronaviridae* family. For the candidate antigens, we chose the C-terminal domain (CTD) of the HCoV-OC43 spike protein in addition to the RBD of the MERS-CoV spike discussed above. The CTD of the HCoV-OC43 spike structurally corresponds to the RBD of the MERS-CoV spike, but has not traditionally been called an RBD because no receptor has been identified for this domain. To reduce the time and effort required for the establishment of a stable cell line, we cotransfected five different plasmids [two HDR donor plasmids (one encoding the MERS antigen and the other encoding the OC43 antigen), two sgRNA-encoding plasmids (targeting the SEMA6A and DCC loci), and a Cas9-encoding plasmid] into HEK-293 cells to simultaneously integrate the MERS-RBD-Fc and OC43-CTD-Fc genes into the SEMA6A and DCC loci, respectively (**Figure 4**a and Figure S16, Supporting Information). Because the two HDR donor plasmids have different homology arm sequences, each gene should integrate into the corresponding hotspot region.

After selection with puromycin and zeocin, surviving colonies were isolated to select cell lines expressing both antigens efficiently. We also wished to compare antigen expression and immunogenicity when the antigens were expressed separately. Therefore, in parallel, we additionally established cell lines that express a single antigen, either MERS-RBD-Fc or OC43-CTD-Fc, by transfecting a single donor plasmid with the corresponding sgRNA-encoding plasmid. We ultimately obtained three different single cell-derived clones that express the MERS antigen only (#1), the human coronavirus OC43 antigen only (#2), and
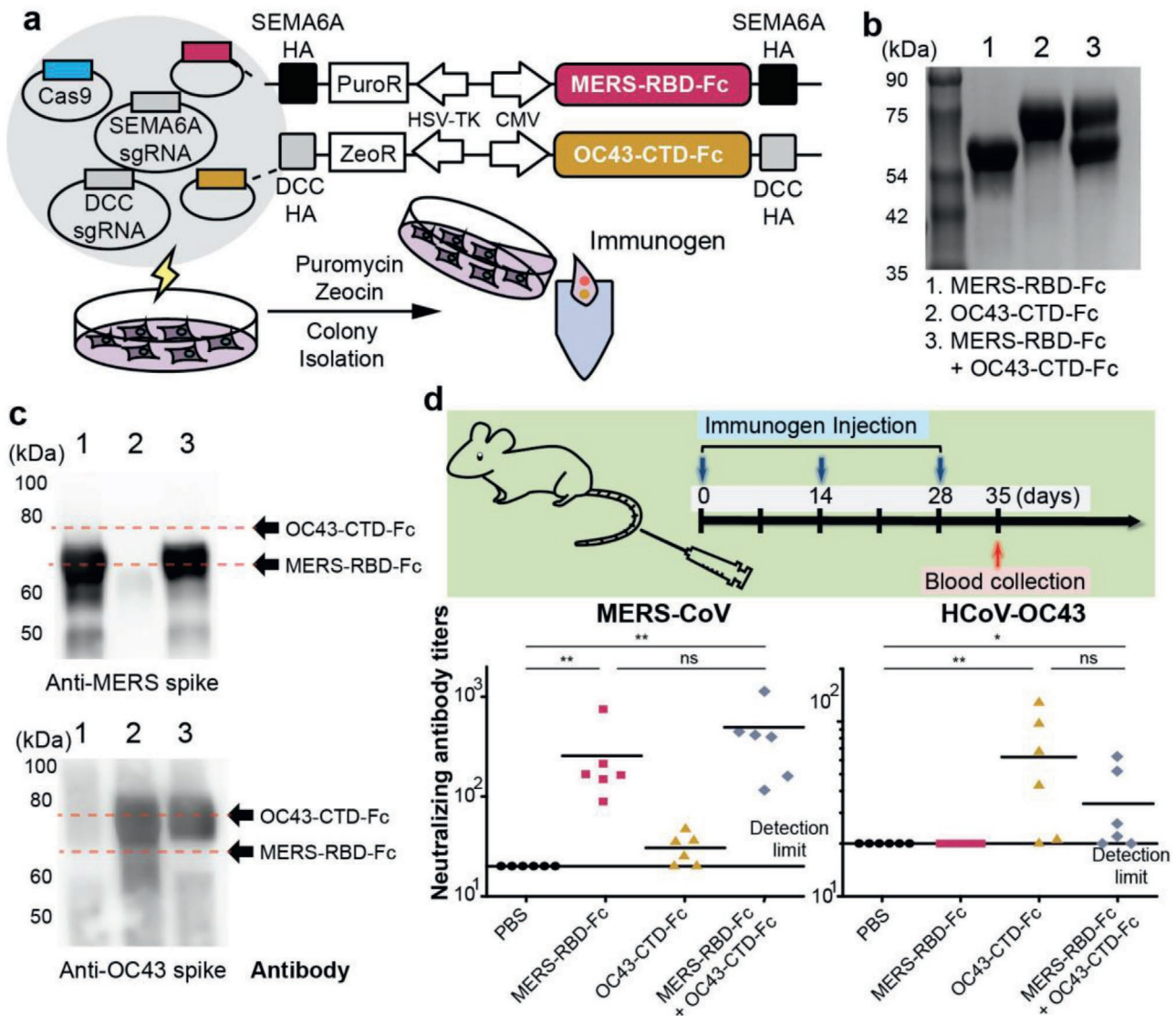
**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
BIOLOGY**

www.advanced-bio.com

Figure 4. Neutralizing activity of two antigens simultaneously expressed in a single cell. a) Top: Schematic representation of vectors expressing the MERS antigen and the human coronavirus OC43 antigen. Bottom: Generation of stable cell lines with multiplexed antigen expression. Knock-in of the sequences expressing the two antigens was done simultaneously. b) Coomassie-stained SDS-PAGE gel of purified proteins from cell lines expressing the MERS antigen (#1), OC43 antigen (#2), or both antigens (#3). c) Purified protein identities were further confirmed by Western blot analysis using an anti-MERS spike antibody and an anti-OC43 spike antibody. Dotted red lines indicate the positions of the MERS and OC43 antigens. d) Top: Mouse immunization and blood collection schedule. Bottom: Neutralizing activity of sera derived from immunized mice against MERS-CoV infection (left-hand panel) and human coronavirus OC43 infection (right-hand panel). X-axis, antigens used for immunization; Y-axis, neutralizing antibody titers (mean, $n = 6$). p-values were analyzed with one-way ANOVA (ns, $p > 0.1$; *$p \approx 0.1$; **$p < 0.05$).

both antigens together (#3), and cultured them in CELLine bioreactor flasks for one week. We purified the antigens from the harvested media using protein G resin and confirmed the identities of the purified antigens by Western blotting with anti-MERS spike and anti-OC43 spike antibodies (Figure 4b). The final protein yields for MERS antigen only, OC43 antigen only, and divalent MERS and OC43 antigen cultures were 7.3 mg, 6.2 mg, and 9.6 mg per liter culture, respectively. Notably, the purified antigen sample from the coexpression cell line contained the two antigens in comparable amounts.

To investigate the immunogenicity of the MERS, OC43, and divalent MERS and OC43 antigens, six mice per antigen

were immunized three times at two-week intervals by injecting purified antigens with alum adjuvants (Figure 4d). We collected mouse sera samples at 35 d after the first injection, and tested their protective effects against viral infection in cultured cells, which is an important indicator of the neutralizing activity of antibodies in sera. Based on the cell viability after viral infection, the sera from mice immunized with the MERS antigen appeared to effectively neutralize MERS-CoV infectivity, but could not neutralize HCoV-OC43 infection at all (Figure 4d). Similarly, the sera from mice immunized with the OC43 antigen showed a neutralizing activity against HCoV-OC43 infectivity but only a little activity against MERS-CoV

(Figure 4d). These results suggest that the antibodies produced following MERS immunization do not cross-react with HCoV-OC43 and vice versa. Remarkably, the sera from mice immunized with the divalent antigen sample showed potent neutralizing activities against both MERS-CoV and HCoV-OC43 viruses, suggesting that the coexpressed and copurified antigens are properly folded and functional (Figure 4d).

3. Discussion

Recombinant proteins produced in human cells are extremely useful for biomedical studies of human or human-virus proteins that require human-specific lipid environments and/or PTMs for functional expression. To obtain such purified proteins, researchers typically transfect DNA plasmids containing the coding sequences of interest into human cells to induce transient overexpression. However, transient expression systems do not always provide a high level of expression and are difficult to scale up. Furthermore, because transfection efficiency is highly variable cell-to-cell, transient expression systems have limited applicability to the coproduction of multiple target recombinant proteins. Alternatively, a stable expression cell line can be generated by randomly integrating target genes into a host chromosome. Although such a system can easily be scaled up, there is no guarantee that the cell line will express the target gene at a high level, and the cell selection process is time-consuming and labor-intensive. In addition, it is extremely difficult to establish single cell-derived clones that simultaneously produce multiple proteins at controllable high levels.

In this study, we demonstrated a stable system for coexpressing multiple proteins in human cells via targeted knock-in of the desired genes into transcriptional hotspots, as a proof-of-concept. We first verified high-level transcription of the EYFP gene from seven selected hotspots and found that EYFP is expressed at similar levels from all tested loci. Additionally, as more EYFP gene copies were sequentially integrated into different hotspots, the EYFP expression level increased gradually with the gene copy number. These findings led us to further develop this useful method for coexpressing multiple different genes or increasing the expression of a single gene. Importantly, because targeted knock-in occurs at desired sites, we avoid problems associated with random integration, which frequently affects the cellular phenotype by causing the loss or gain of endogenous gene function. Knock-in at specific sites is both safer and more predictable. Taken together, we envision that our method can be used for the production of multi-subunit protein complexes and cytotoxic proteins together with inhibitors, which have been hardly produced in human cell system. We also like to emphasize that our strategy can generally be applied for other expression system from different organisms such as CHO cells. If transcription hotspots of the CHO cells are determined, one could develop similar coexpression method in CHO cells through the targeted integration of the desired genes into the sites.

Furthermore, we constructed a single cell-derived clone that expressed both MERS-CoV and OC43 virus antigens and showed that antibodies raised against the two antigens could neutralize an infection of either of the two viruses in vitro. Human coronaviruses are pathogenic, causing acute respiratory

diseases in humans. Since 2002, there have been pandemics of severe acute respiratory syndrome coronavirus, MERS-CoV, and the recent 2019 novel coronavirus (COVID-19),[41] all of which can cause deadly pneumonia. Four additional human coronaviruses (HCoV 229E, NL63, OC43, HKU1) are circulating in the human population causing mild symptoms.[42] Currently, no effective therapeutics or vaccines have been reported for any of these coronaviruses. Therefore, therapeutic and preventive strategies are urgently needed. In the case of COVID-19, our system could be used to produce the type of recombinant subunit vaccines (divalent and carrying appropriate PTMs) that are most effective in humans. Alternatively, the system could be used to produce the RBD of the COVID-19 spike protein, which could be purified and used to develop inhibitor drugs that would prevent the interaction of the virus with its known receptor, to detect neutralizing antibodies from patients recovering from infection, and to identify novel human receptors. Taken together, results from our study suggest a novel means of producing any human-related protein, including viral antigens, in the most effective and inexpensive way.

4. Experimental Section

Cells, Viruses, and Antibodies: MERS-CoV was provided by the Korea Centers for Disease Control and Prevention (1-001-MER-CO-2015001). HCoV-OC43 (VR-1558), Vero cells (CRL-1586), RD cells (CCL-136), and HEK-293 cells (CRL1573) were purchased from the American Type Culture Collection (ATCC). Huh-7 cells were obtained from the Japan Cell Research Bank. MERS-CoV was expanded using Huh-7 cells and the viral titer was measured by a plaque assay. HCoV-OC43 was expanded using RD cells and the viral titer was determined by a $TCID_{50}$ assay. HRP-conjugated antihuman IgG antibody, used to detect the human Fc protein, was purchased from Thermo Fisher Scientific (#31410). Antispike$^{MERS-CoV}$ antibody was obtained from Sino Biological (#40069-RP02). Anti-OC43 spike protein antibody was purchased from CUSABIO (#CSB-PA336163EA01HIY).

Cell Culture Conditions: T-REx-293 cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 μg mL^{-1} streptomycin, 100 units mL^{-1} penicillin, 0.1 × 10^{-3} M nonessential amino acids, and 5 μg mL^{-1} blasticidin. Vero cells, RD cells, and HEK-293 cells were maintained in DMEM supplemented with 10% FBS. For antibiotic selection, cells were maintained in the presence of hygromycin B (Thermo Fisher, 10687010) at 200 μg mL^{-1} for at least 14 d, puromycin (InvivoGen, ant-pr-1) at 1 μg mL^{-1} for at least 7 d, or Zeocin (InvivoGen, ant-zn-05) at 200 μg mL^{-1} for at least 7 d.

Transfection Conditions: 1.0×10^5 T-REx-293 cells and HEK-293 cells were plated at 1 d before transfection. Mixture of SpCas9 expression plasmid (750 ng), sgRNA-encoding plasmids (250 ng), and donor templates (500 ng) or mixture of LbCas12a expression plasmid (800 ng), crRNA-encoding plasmids (200 ng) and donor templates (500 ng) were transfected into cells using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Genomic DNA was isolated at 72 h after transfection.

Targeted Deep Sequencing: T-REx-293 cells, which had been cultivated for 3 d, were pelleted by centrifugation. The cell pellet was mixed with 100 μl of Proteinase K extraction buffer [40 × 10^{-3} M Tris-HCl (pH 8.0), 1% Tween-20, 0.2 × 10^{-3} M EDTA, 10 mg of Proteinase K, 0.2% nonidet P-40] and incubated at 60 °C for 15 min followed by an incubation at 98 °C for 5 min. The extracted genomic DNAs were amplified using SUN-PCR blend (SUN GENETICS), which was confirmed by gel electrophoresis methods such as a knock-in specific PCR. The resulting PCR products were analyzed using an Illumina Mini-Seq instrument. The results of Mini-seq were analyzed using Cas-Analyzer (http://www.rgenome.net/be-analyzer/).[37]

**ADVANCED
SCIENCE NEWS**

www.advancedsciencenews.com

**ADVANCED
BIOLOGY**

www.advanced-bio.com

Flow Cytometry Analysis: To induce expression of the gene of interest in the donor template, T-RExTM-293 cells were treated with doxycycline hyclate (Sigma-Aldrich, D9891) for 1 d before fluorescence-activated cell sorting (FACS) analysis using FACSAriaIII and BD FACSCanto II instruments. The results were analyzed using FlowJo software.

Construction of Antigen-Encoding Donor Plasmids: The MERS-CoV antigen-encoding sequence was constructed by fusion of sequences encoding the signal sequence (residues 1–18) and RBD (residues 367–606) of the MERS-CoV spike protein and the human IgG Fc fragment. The HCoV-OC43 antigen-encoding sequence was constructed by fusion of sequences encoding the signal sequence of the MERS-CoV spike protein, the CTD of the HCoV-OC43 spike protein (residues 324–685), and the human IgG Fc fragment. The coding sequences of the two antigen constructs were synthesized by Integrated DNA Technologies. The coding sequence of the MERS-CoV antigen was inserted into an HDR donor plasmid for integration into the AAVS1 or SEMA6A locus, whereas that of the HCoV-OC43 antigen was inserted into an HDR donor plasmid for integration into the DCC locus.

Generation of Antigen-Expressing Cell Lines: HEK-293 cells (ATCC CRL-1573) were cultured in DMEM supplemented with 10% FBS and transfected with a donor plasmid containing sequences encoding the MERS-CoV or HCoV-OC43 antigen, an sgRNA expression plasmid, and the SpCas9 expression plasmid. The transfected cells were cultured in medium containing 200 µg mL^{-1} hygromycin B, 1 µg mL^{-1} puromycin, and/or 200 µg mL^{-1} Zeocin for 14 d, and surviving colonies were isolated. A single colony was picked up by pipet and then resuspended into selection media in 96-well plate. Isolated colonies were grown and frozen for further experiment. To determine the level of antigen expression, the culture medium was subjected to Western blot analysis. Because most colonies showed a similar high level of expression, one of the high expression cell lines was selected for large-scale production.

Large-Scale Antigen Expression and Purification: Antigen-expressing cells grown to confluence in 12 of 150 mm cell culture plates were harvested, resuspended in 30 mL FreeStyle 293 medium (ThermoFisher Scientific, #12338018), and seeded into the lower chamber of a CELLine bioreactor flask (Sigma-Aldrich, #Z688029). The upper chamber of the flask was filled with 1 L FreeStyle 293 medium. After 4 d, 4 × 10^{-3} M valproic acid (Sigma-Aldrich, #p4543) was added to the upper chamber to increase protein expression. In 3 d after valproic acid treatment, the culture supernatant in the lower chamber was harvested, filtered using a 0.22 µm membrane filter, and loaded onto a protein G column (GE Healthcare, #17-0618-01). The bound proteins were eluted with 100 × 10^{-3} M glycine buffer (pH 2.7). The eluted sample was neutralized with 1 M Tris-HCl (pH 9.0) using 30% volume of elution buffer and dialyzed against the phosphate-buffered saline (PBS) and concentrated using Amicon Ultra-15 centrifugal filter (Millipore, #UFC901024) and analyzed by SDS-PAGE.

N-deglycosylation Assay: The purified MERS-CoV antigen was treated with PNGase F (New England Biolabs, #P0704S) according to the manufacturer's instructions. For this assay, MERS antigen fused only with histidine tag without human Fc was produced with similar method. The mobility shift of the antigen on an SDS-PAGE gel was identified by Western blot using a 6-His Tag antibody (Bethyl Laboratories, #A190-114p).

Mouse Immunization: Animal experiments were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the Korea Research Institute of Chemical Technology (2019-8B-08-01). Twenty four six-weeks-old female C57BL/6 mice were evenly divided into four groups. Each group was immunized with PBS, the MERS-CoV antigen, the HCoV-OC43 antigen, or the divalent antigen (MERS + OC43) sample. 100 µg of antigen with 1 mg Alum adjuvant (InvivoGen, #vac-alu-250) was injected intramuscularly into each mouse three times at two-week intervals. At 35 d after the first injection, mouse serum samples were collected to test their neutralizing activity against viral infection.

Neutralization Assay: Microneutralization assays were performed to determine the neutralizing activity of the serum samples. Mouse serum was serially diluted by 2-folds with minimum essential media (MEM) supplemented with 2% FBS, from 1/20 to 1/1280, in 96-well plates. Each diluted serum sample was mixed with MERS-CoV or HCoV-OC43 and incubated for 1 h at 37 °C. To test the neutralization activity against MERS-CoV, Vero cells were overlaid with a mixture of mouse serum and MERS-CoV. After 3 d of incubation, cell viability was measured by a colorimetric assay using MTS solution (Promega, #G3582) and a Synergy H1 microplate reader (BioTek) according to the manufacturers' instructions. To test the neutralization activity against HCoV-OC43, RD cells were overlaid with a mixture of mouse serum and HCoV-OC43. After 3 d of incubation, cell viability was measured by the MTT assay.[43] The neutralizing antibody titers were calculated as the reciprocal of the highest dilution of sera that protected 50% of the cells. Mock-infected cells were considered to show 100% survival, and virus-infected cells were considered to show 0% survival.

Statistical Analysis: Statistical analyses were performed using one-way ANOVA.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Notes

C.K., Y.K.J., J.Y., S.-J.K., B.-T.K., H.J.S., K.B.K., J.-S.W., and S.B. have filed two patent applications based on this work.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

C.K. and Y.K.J. contributed equally to this work. C.K., S.-J.K., J.-S.W., and S.B. conceived this project; Y.K.J., J.Y., H.J.C., J.H.H., and S.-A.H. performed experiments to reveal the characteristics of transcriptional hotspots and C.K., H.J.S., and K.B.K. performed experiments for vaccine production and assessing the neutralizing activities; S.-J.K. gave critical comments; C.K., Y.K.J., J.-S.W., and S.B. wrote the manuscript with the approval of all other authors. S.-J.K., B.-T.K., J.-S.W., and S.B. supervised the research.

Data Availability Statement

High–throughput sequencing data have been deposited in the NCBI Sequence Read Archive database (SRA; https://www.ncbi.nlm.nih.gov/sra) under accession number PRJNA612584.

[1] D. Gray, *Curr. Protoc. Protein Sci.* **1997**, *10*, 591.

[2] S. M. Noh, M. Sathyamurthy, G. M. Lee, *Curr. Opin. Chem. Eng.* **2013**, *2*, 391.

[3] A. Croset, L. Delafosse, J.-P. Gaudry, C. Arod, L. Glez, C. Losberger, D. Begue, A. Krstanovic, F. Robert, F. Vilbois, L. Chevalet, B. Antonsson, *J. Biotechnol.* **2012**, *161*, 336.

[4] E. Lorenzo, L. Méndez, E. Rodríguez, N. Gonzalez, G. Cabrera, C. Pérez, R. Pimentel, Y. Sordo, M. P. Molto, T. Sardina, A. Rodríguez-Mallon, M. P. Estrada, *AMB Express* **2019**, *9*, 139.

[5] K.-J. Kim, H.-E. Kim, K.-H. Lee, W. Han, M.-J. Yi, J. Jeong, B.-H. Oh, *Protein Sci.* **2004**, *13*, 1698.

[6] H. Jonckheere, K. De Vreese, Z. Debyser, J. Vandekerckhove, J. Balzarini, J. Desmyter, E. De Clercq, J. Anné, *J. Virol. Methods* **1996**, *61*, 113.

[7] I. Berger, D. J. Fitzgerald, T. J. Richmond, *Nat. Biotechnol.* **2004**, *22*, 1583.

[8] M. Hansson, P.-Å. Nygren, S. Ståhl, *Biotechnol. Appl. Biochem.* **2000**, *32*, 95.

[9] N. K. Tripathi, A. Shrivastava, *Front. Immunol.* **2018**, *9*, 1919.

[10] L. Du, G. Zhao, Z. Kou, C. Ma, S. Sun, V. K. M. Poon, L. Lu, L. Wang, A. K. Debnath, B.-J. Zheng, Y. Zhou, S. Jiang, *J. Virol.* **2013**, *87*, 11963.

[11] N. J. Lennemann, B. A. Rhein, E. Ndungo, K. Chandran, X. Qiu, W. Maury, *MBio* **2014**, *5*, e00862.

[12] S. Nallet, M. Amacker, N. Westerfeld, L. Baldi, I. König, D. L. Hacker, C. Zaborosch, R. Zurbriggen, F. M. Wurm, *Vaccine* **2009**, *27*, 6415.

[13] M. Suárez, Y. Sordo, Y. Prieto, M. P. Rodríguez, L. Méndez, E. M. Rodríguez, A. Rodríguez-Mallon, E. Lorenzo, E. Santana, N. González, P. Naranjo, M. T. Frías, Y. Carpio, M. P. Estrada, *Vaccine* **2017**, *35*, 4437.

[14] R. Barrangou, C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D. A. Romero, P. Horvath, *Science* **2007**, *315*, 1709.

[15] C. Anders, O. Niewoehner, A. Duerst, M. Jinek, *Nature* **2014**, *513*, 569.

[16] S. W. Cho, S. Kim, J. M. Kim, J.-S. Kim, *Nat. Biotechnol.* **2013**, *31*, 230.

[17] S. H. Lee, S. Bae, *BMB Rep.* **2016**, *49*, 201.

[18] H. Nishimasu, F. A. Ran, P. D. Hsu, S. Konermann, S. I. Shehata, N. Dohmae, R. Ishitani, F. Zhang, O. Nureki, *Cell* **2014**, *156*, 935.

[19] J. M. Kim, D. Kim, S. Kim, J.-S. Kim, *Nat. Commun.* **2014**, *5*, 3157.

[20] H. Kim, J.-S. Kim, *Nat. Rev. Genet.* **2014**, *15*, 321.

[21] J. K. Cheng, A. M. Lewis, D. S. Kim, T. Dyess, H. S. Alper, *Biotechnol. J.* **2016**, *11*, 1100.

[22] R. M. Kotin, R. M. Linden, K. I. Berns, *EMBO J.* **1992**, *11*, 5071.

[23] I. Tan, C. H. Ng, L. Lim, T. Leung, *J. Biol. Chem.* **2001**, *276*, 21209.

[24] R. C. DeKelver, V. M. Choi, E. A. Moehle, D. E. Paschon, D. Hockemeyer, S. H. Meijsing, Y. Sancak, X. Cui, E. J. Steine, J. C. Miller, P. Tam, V. V. Bartsevich, X. Meng, I. Rupniewski, S. M. Gopalan, H. C. Sun, K. J. Pitz, J. M. Rock, L. Zhang, G. D. Davis, E. J. Rebar, I. M. Cheeseman, K. R. Yamamoto, D. M. Sabatini, R. Jaenisch, P. D. Gregory, F. D. Urnov, *Genome Res.* **2010**, *20*, 1133.

[25] K. Suzuki, Y. Tsunekawa, R. Hernandez-Benitez, J. Wu, J. Zhu, E. J. Kim, F. Hatanaka, M. Yamamoto, T. Araoka, Z. Li, M. Kurita, T. Hishida, M. Li, E. Aizawa, S. Guo, S. Chen, A. Goebl, R. D. Soligalla, J. Qu, T. Jiang, X. Fu, M. Jafari, C. R. Esteban, W. T. Berggren, J. Lajara, E. Nuñez-Delicado, P. Guillen, J. M. Campistol, F. Matsuzaki, G.-H. Liu, P. Magistretti, K. Zhang, E. M. Callaway, K. Zhang, J. C. I. Belmonte, *Nature* **2016**, *540*, 144.

[26] A. Lombardo, D. Cesana, P. Genovese, B. Di Stefano, E. Provasi, D. F. Colombo, M. Neri, Z. Magnani, A. Cantore, P. Lo Riso, M. Damo, O. M. Pello, M. C. Holmes, P. D. Gregory, A. Gritti, V. Broccoli, C. Bonini, L. Naldini, *Nat. Methods* **2011**, *8*, 861.

[27] P. Genovese, G. Schiroli, G. Escobar, T. Di Tomaso, C. Firrito, A. Calabria, D. Moi, R. Mazzieri, C. Bonini, M. C. Holmes, P. D. Gregory, M. Van Der Burg, B. Gentner, E. Montini, A. Lombardo, L. Naldini, *Nature* **2014**, *510*, 235.

[28] S. Bae, J. Park, J.-S. Kim, *Bioinformatics* **2014**, *30*, 1473.

[29] J. Park, S. Bae, J.-S. Kim, *Bioinformatics* **2015**, *31*, 1093.

[30] J. C. Miller, M. C. Holmes, J. Wang, D. Y. Guschin, Y. L. Lee, I. Rupniewski, C. M. Beausejour, A. J. Waite, N. S. Wang, K. A. Kim, P. D. Gregory, C. O. Pabo, E. J. Rebar, *Nat. Biotechnol.* **2007**, *25*, 778.

[31] M. Christian, T. Cermak, E. L. Doyle, C. Schmidt, F. Zhang, A. Hummel, A. J. Bogdanove, D. F. Voytas, *Genetics* **2010**, *186*, 757.

[32] T. Gaj, C. A. Gersbach, C. F. Barbas, *Trends Biotechnol.* **2013**, *31*, 397.

[33] B. Zetsche, J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin, F. Zhang, *Cell* **2015**, *163*, 759.

[34] H. K. Kim, M. Song, J. Lee, A. V. Menon, S. Jung, Y.-M. Kang, J. W. Choi, E. Woo, H. C. Koh, J.-W. Nam, H. Kim, *Nat. Methods* **2017**, *14*, 153.

[35] H. K. Kim, S. Min, M. Song, S. Jung, J. W. Choi, Y. Kim, S. Lee, S. Yoon, H. Kim, *Nat. Biotechnol.* **2018**, *36*, 239.

[36] J. Jones, T. Nivitchanyong, C. Giblin, V. Ciccarone, D. Judd, S. Gorfien, S. S. Krag, M. J. Betenbaugh, *Biotechnol. Bioeng.* **2005**, *91*, 722.

[37] J. Park, K. Lim, J.-S. Kim, S. Bae, A. Valencia, *Bioinformatics* **2017**, *33*, 286.

[38] Y. Wang, W. Tai, J. Yang, G. Zhao, S. Sun, C. T. K. Tseng, S. Jiang, Y. Zhou, L. Du, J. Gao, *Hum. Vaccines Immunother.* **2017**, *13*, 1615.

[39] H. Garg, T. Mehmetoglu-Gurbuz, A. Joshi, *Sci. Rep.* **2020**, *10*, 4017.

[40] R. Yaesoubi, C. Trotter, C. Colijn, M. Yaesoubi, A. Colombini, S. Resch, P. A. Kristiansen, F. M. LaForce, T. Cohen, *PLoS Med.* **2018**, *15*, e1002495.

[41] E. Mahase, *BMJ* **2020**, *368*, m641.

[42] X. Ou, H. Guan, B. Qin, Z. Mu, J. A. Wojdyla, M. Wang, S. R. Dominguez, Z. Qian, S. Cui, *Nat. Commun.* **2017**, *8*, 15216.

[43] H. Kanga, C. Kim, D.-e. Kim, J. H. Song, M. Choi, K. Choi, M. Kang, K. Lee, H. S. Kim, J. S. Shin, J. Kim, S.-B. Han, M.-Y. Lee, S. U. Lee, C. K. Lee, M. Kim, H.-J. Ko, F. J. M. van Kuppeveld, S. Cho, *Antiviral Res.* **2015**, *124*, 1.